arXiv:2304.00707v2 [math.PR] 3 Apr 2024

具有光滑协方差的在线最小二乘随机梯度下降的高维尺度极限与涨落

Krishnakumar Balasubramanian* Promit Ghosal † Ye He ‡

摘要

我们在明确考虑数据生成模型性质的基础上，推导了在线最小二乘随机梯度下降（SGD）算法的高维尺度极限与涨落。我们的方法将随机梯度下降迭代视为一个相互作用粒子系统，其期望相互作用由输入的协方差结构刻画。在假设最高至八阶矩满足光滑性条件，且不显式假设高斯性的情况下，我们建立了以无限维常微分方程（ODE）或随机微分方程（SDE）形式表示的高维尺度极限与涨落。结果揭示了迭代过程精确的三阶段相变：随着噪声方差从低水平升至中等水平，最终达到极高水平，迭代过程依次呈现弹道态、扩散型和纯随机行为。在低噪声情形下，我们进一步以无限维随机微分方程刻画了（缩放后）迭代过程的精确涨落。我们还证明了所推导的极限常微分方程和随机微分方程的解存在且唯一。我们的结果有多种应用，包括刻画极限均方估计误差或预测误差及其涨落；这些量可通过解析或数值求解极限方程得到。

目录

1 引言 **2**

1.1 预备知识.3 1.2 假设.5 1.3 主要结果.8 1.4 相关工作.11

2 应用与示例 **13**

2.1 均方误差（MSE）和预测误差（PE）的尺度极限与涨落.14 2.2 补充见解与具体示例.15

3 证明概述 **17**

3.1 定理 1.6 的证明.17 3.2 定理 1.10 的证明.18 3.3 解的存在性与唯一性.20

*加州大学戴维斯分校统计系。该作者的研究部分得到了美国国家科学基金会（NSF）DMS-2053918 项目的资助。kbala@ucdavis.edu

†布兰代斯大学数学系。该作者的研究部分得到了美国国家科学基金会（NSF）DMS-2153661 项目的资助。promit@brandeis.edu

‡佐治亚理工学院数学学院。该作者的研究部分得到了授予加州大学戴维斯分校的美国国家科学基金会 TRIPODS 项目 CCF-1934568 的资助。leohe@ucdavis.edu

4 尺度极限的证明 **20**
4.1 $\{\Theta^{d,T}\}_{d\geq 1,\ T>0}$的紧性.20 4.2 极限识别.27

5 涨落的证明 **33**
5.1 $\{U^{d,T}\}_{d\geq 1,}$ 的紧性.34 5.2 极限识别.39

6 随机微分方程存在性与唯一性的证明 48

7 第 2 节应用的证明 54

1 引言

随机梯度下降（SGD）最初是由 RobbinsandMonro （1951年）提出的，用于求根。如今，SGD及其变体已成为训练机器学习模型最常用的算法，应用范围从大规模线性模型到深度神经网络。理解SGD所面临的主要挑战之一，是弄清其收敛性质。对于固定维度的问题，过去二十年来，学习理论和优化领域一直致力于给出非渐近界，既包括期望意义下的界，也包括高概率界。然而，在实践中出现的大规模统计问题通常都基于特定的数据生成模型，而这类界在预测SGD算法在这些问题上的实际行为时，往往过于保守。为了解决这一问题，近期研究开始致力于刻画大规模高维问题中SGD的精确动力学。具体而言，研究重点是在适当的缩放条件下，当迭代次数或观测值数量以及数据维度趋于无穷时，获得SGD精确的渐近行为及其涨落。这一方法背后的主要思想是证明，在所考虑的缩放条件下，SGD中的噪声效应会相互平均抵消，因此其精确的渐近行为和涨落由一组特定的动力系统方程决定。

本文旨在研究特定统计问题（即线性回归问题）中的随机梯度下降算法，并对其在高维缩放下的行为进行细致分析。具体而言，我们考虑线性回归模型 $Y = X^\top \theta^* + \mathcal{E}$，其中 $\theta^* \in \mathbb{R}^d$ 是真实回归系数， $X \in \mathbb{R}^d$ 是均值为零、协方差矩阵为 $\mathbb{E}[XX^\top] = \Sigma_d \in \mathbb{R}^{d\times d}$ 的输入随机向量， $Y \in \mathbb{R}$ 是输出或响应随机变量，而 $\mathcal{E} \in \mathbb{R}$ 是至少具有有限方差的零均值噪声。对于该统计模型，我们考虑求解以下总体最小二乘随机优化问题

$$\min_{\theta\in\mathbb{R}^d} \mathbb{E}[(Y - \langle X, \theta\rangle)^2],$$

使用初始猜测值为 $\theta^0 \in \mathbb{R}^d$ 的在线随机梯度下降，其迭代过程如下

$$\theta^{t+1} = \theta^t + \eta \left(y^t - \langle x^t, \theta^t\rangle\right) x^t, \tag{1}$$

其中， $\theta^t \in \mathbb{R}^d$ 是时刻 t的输出。序列 $\{x^t, y^t\}_{t\geq 0}$ 对应观测值，其中 $y^t = \langle x^t,\ \theta^*\rangle + \varepsilon^t$ 且 $\eta > 0$ 是步长参数，在获得我们的尺度极限和涨落方面起着关键作用。关于 η 的具体取值将在第1.3节中详细说明。序列 $\{x^t\}_{t\geq 0}$ 和 $\{\varepsilon^t\}_{t\geq 0}$ 分别被假设为随机向量 X 和噪声 $\mathcal{E}$ 的独立同分布副本。特别注意，对于式（1）中的在线随机梯度下降，迭代次数等于所使用的观测值数量。

在我们的分析中，我们将式（1）中的最小二乘在线随机梯度下降视为一个离散的时空相互作用粒子系统，其中空间轴对应向量 θ^t 的各个坐标，时间轴对应算法的演化。具体而言，注意，式（1）中的在线随机梯度下降更新可表示为以下逐坐标形式。对于任意 $1 \le i \le d$,

$$\begin{aligned}\theta_i^{t+1} &= \theta_i^t + \eta x_i^t y^t - \eta \sum_{j=1}^d x_i^t x_j^t \theta_j^t \\ &= \theta_i^t + \eta x_i^t \Big(\sum_{j=1}^d x_j^t \theta_j^* + \varepsilon^t - \sum_{j=1}^d x_j^t \theta_j^t \Big) \\ &= \theta_i^t - \eta \sum_{j=1}^d x_i^t x_j^t (\theta_j^t - \theta_j^*) + \eta x_i^t \varepsilon^t.\end{aligned}$$

现在，将集中化迭代量定义为 $\Delta\theta^t := \theta^t - \theta^*$，并令 $\Delta\theta_i^t$ 表示其第 i 个坐标，对所有 $1 \le i \le d$ 和 $t \ge 0$ 均成立，则最小二乘在线随机梯度下降也可表示为以下相互作用粒子系统:

$$\Delta\theta_i^{t+1} = \Delta\theta_i^t - \underbrace{\eta \sum_{j=1}^d x_i^t x_j^t \Delta\theta_j^t}_{random\ \text{interaction}} + \eta x_i^t \varepsilon^t, \qquad 1 \le i \le d, \tag{2}$$

其中，粒子 $\{\Delta\theta_i^t\}_{1\le i\le d}$ 彼此相互作用，并随离散时间尺度演化。具体而言，对于任意 t，粒子 $\{\Delta\theta_i^t\}_{1\le i\le d}$ 之间的相互作用是随机的，而期望相互作用由输入向量 X的 Σ_d 协方差矩阵刻画。因此，为了分析最小二乘在线随机梯度下降的高维渐近性质，我们分析式（2）给出的相互作用粒子系统的标度极限与涨落。具体而言，我们的极限以无穷维常微分方程（常微分方程）和随机微分方程（随机微分方程）的形式推导；关于相关背景，参见 Arnold (1992年); Kallianpur 和 Xiong (1995);Da Prato 和 Zabczyk (2014年).

我们的方法也受到分析相互作用粒子系统的大量相关文献的启发。例如，参见Kipnis 和 Landim(1998年); Liggett(1999年);Zeitouni(2004);Sznitman(2004); Darling 和 Norris (2008年); Spohn (2012 年)。式（ 2 ）中的相互作用粒子系统可表现为长程相互作用或短程相互作用，具体取决于协方差矩阵的结构Σ_d。当协方差矩阵 Σ_d 在适当意义上是“平滑”的，从而避免协方差矩阵各元素发生突变时，对应的就是长程相互作用情形。这类协方差矩阵的例子包括实践中常见的可带状化协方差矩阵和循环协方差矩阵（Gray 等，2006）。

1.1 预备知识

在继续之前，我们先列出本文中使用的记号。对于正整数 a，我们令$[a] := \{1, \dots, a\}$。对于向量，上标表示时间索引，下标表示坐标。 S（欧几里得空间的一个子集）上平方可积函数构成的空间记为 $L^2(S)$，其平方范数为 $\|g\|_{L^2(S)}^2 := \int_S g(x)^2 dx$ 对任意 $g \in L^2(S)$ 均成立。 $[0,1]$, 上的连续函数空间记为 $C([0,1])$，并赋予关于$[0,1]$的一致收敛拓扑。 $[0,1]$, 上具有连续导数直至 k^{th} 阶的连续函数空间记为 $C^k([0,1])$，并赋予函数值上一致收敛的拓扑

以及关于 $[0,1]$ 上直至 k^{th} 阶导数的一致收敛拓扑（我们主要使用 $k=1,2$）。对于任意拓扑空间 $\mathcal{H}$ 和 $\tau>0$，$C([0,\tau];\mathcal{H})$ 表示具有连续轨迹的 $\mathcal{H}$ 值函数空间。

我们还需要定义高斯随机场或高斯随机场过程，它们用于刻画随机梯度下降迭代的极限行为。更多背景信息请参阅 Adler and Taylor(2007。

定义 1.1。 高斯随机场 是定义在参数集 $[0,1]$ 上的随机场 g，对于每个 $1\le K<\infty$ 以及每个 $(x_1,\cdots,x_K)\in[0,1]^K$，其有限维分布 $(g(x_1)\cdots,g(x_K))$ 均为多元高斯分布。高斯随机场过程 是定义在参数集 $[0,\infty)$ 上的时间索引随机场 g，其中 $\times[0,1]$，对于每个 $1\le K<\infty$ 以及每个 $(x_1,\cdots,x_K)\in[0,1]^K$，$\{(g(t,x_1)\cdots,g(t,x_K))\}_{t\ge0}$ 均为多元高斯过程。

时空插值。 我们的方法首先通过对（2）中的离散粒子进行分段线性插值，构造一个时空随机过程。该过程记为 $\{\Theta^{d,T}(s,\ x)\}_{s\in[0,\tau],\ x\in[0,1]}$，在时间和空间上均连续。宏观尺度上的空间坐标由集合 $[0,1]$ 索引，记为空间变量 x。微观尺度上的空间分辨率为 $1/d$量级。令 $T\in\mathbb{N}_+$ 为一个正整数参数。参数 $\tau\in(0,\ \infty)$ 和 $1/T$ 分别对应于我们希望观测轨迹的时间尺度以及时间轴的分辨率。对于任意 τ，我们考虑前 $\lfloor\tau T\rfloor$ 个最小二乘在线随机梯度下降迭代。具体而言，$\lfloor\tau T\rfloor$ 对应于迭代总次数，也对应于所使用的观测值数量。

现在我们描述如何构造（2）中离散粒子的插值。考虑函数 $\Theta^{d,T}(\cdot,\cdot):[0,\tau]\times[0,1]\to\mathbb{R}$，它满足以下条件：

(a) 在时间变量上以网格宽度为 T^{-1}、空间变量上以网格宽度为 d^{-1} 的网格点处，其取值等于随机梯度下降粒子的位置；也就是说，对任意 $i\in[d]$ 和任意 $0\le t\le\lfloor\tau T\rfloor$，

$$\Theta^{d,T}\Big(\frac{t}{T},\frac{i}{d}\Big)=\Delta\theta_i^t.$$

我们可以人为地定义 $\Delta\theta_0^t=\Delta\theta_1^t$。

(b) 它关于时间变量是分段常数的；也就是说，对任意 $s\in[0,\tau]$，

$$\Theta^{d,T}(s,\cdot)=\Theta^{d,T}\Big(\frac{\lfloor sT\rfloor}{T},\cdot\Big).$$

(c) 它关于空间变量是分段线性的；也就是说，对任意 $x\in[0,1]$，

$$\Theta^{d,T}(\cdot,x)=(\lfloor dx\rfloor+1-dx)\Theta^{d,T}\Big(\cdot,\frac{\lfloor dx\rfloor}{d}\Big)+(dx-\lfloor dx\rfloor)\Theta\Big(\cdot,\frac{\lfloor dx\rfloor+1}{d}\Big).$$

根据上述构造，我们可以基于 $\{\Delta\theta_i^t\}_{i\in[d],0\le t\le\lfloor\tau T\rfloor}$ 显式表示 $\Theta^{d,T}$：对任意$s\in[0,\tau]$ 和任意 $x\in[0,1]$，

$$\begin{aligned}\Theta^{d,T}(s,x)&=(\lfloor dx\rfloor+1-dx)\Theta^{d,T}\Big(\frac{\lfloor sT\rfloor}{T},\frac{\lfloor dx\rfloor}{d}\Big)+(dx-\lfloor dx\rfloor)\Theta\Big(\frac{\lfloor sT\rfloor}{T},\frac{\lfloor dx\rfloor+1}{d}\Big)\\&=(\lfloor dx\rfloor+1-dx)\Delta\theta_{\lfloor dx\rfloor}^{\lfloor sT\rfloor}+(dx-\lfloor dx\rfloor)\Delta\theta_{\lfloor dx\rfloor+1}^{\lfloor sT\rfloor}.\end{aligned}$$

条件 (b) 尤其意味着，在时间变量上，$\Theta^{d,\ T}$ 是 c`adla`g 函数。由条件 (c) 可容易看出，$\Theta^{d,\ T}$ 在空间上连续，且具有定义良好的弱导数。实际上，

$\Theta^{d,\,T}$的一阶空间导数对除网格点 $\{0, \frac{1}{d}, \cdots, \frac{d-1}{d}, 1\}$外的所有 x 均有定义。接下来，我们构造在时间和空间上均为分段线性的 $\Theta^{d,\,T}$。对于任意 $s \in [0, \tau]$ 和 $x \in [0, 1]$,

$$\overline{\Theta}^{d,T}(s,x) \coloneqq (\lfloor sT \rfloor + 1 - sT)\,\Theta^{d,T}\Big(\frac{\lfloor sT \rfloor}{T}, x\Big) + (sT - \lfloor sT \rfloor)\,\Theta^{d,T}\Big(\frac{\lfloor sT \rfloor + 1}{T}, x\Big). \tag{3}$$

由此构造可立即看出 $\Theta^{d,\,T} \in C([0,\tau];\, C[0,1])$。本研究的主要目标是在适当的缩放下，刻画当 T 和 d 均趋于无穷时，时空随机过程$\{\Theta^{d,T}(s,\ x)\}_{s\in[0,\tau],\ x\in[0,1]}$的极限行为；这里步长参数 η取适当的值，并对噪声 $\mathcal{E}$作出各种假设。

1.2 假设

我们首先给出时空随机过程的如下定义，它刻画了 X （或者等价地，序列 $\{x^t\}_{t\geq 0}$）的数据生成过程。

定义 1.2 （数据生成过程）。设 A，B 和 E 是定义在 $[0,1]^2$, $[0,1]^4$ 和 $[0,1]^8$ 上的实值对称函数。$\{W(s,\ x)\}_{s\geq 0,\ x\in[0,1]}$ 是一个在时间变量上为白噪声的随机过程，使得对所有 $s \geq 0$ 和 x，x_1，$\cdots$，$x_8 \in [0,1]$:

$$\mathbb{E}\,[W(s,x)] = 0, \qquad \mathbb{E}\,[W(s,x_1)W(s,x_2)] = A(x_1,x_2),$$
$$\mathbb{E}\big[\prod_{i=1}^{4} W(s,x_i)\big] = B(x_1,x_2,x_3,x_4), \qquad \mathbb{E}\big[\prod_{i=1}^{8} W(s,x_i)\big] = E(x_1,\cdots,x_8).$$

现在我们介绍关于样本$\{x_i^t\}_{t\geq 0、\ i\in[d]}$和随机噪声 $\{\varepsilon^t\}_{t\geq 0}$的假设。

假设 1.1。样本$\{x_i^t\}_{1\leq i\leq d、\ t\geq 0}$ 和噪声 $\{\varepsilon^t\}_{t\geq 0}$ 相互独立，并满足

(a) 存在一个随机过程 $\{W(s,x)\}_{s\geq 0, x\in[0,1]}$，其定义见 *1.2*。

$$x_i^t = W\Big(\frac{t}{T}, \frac{i}{d}\Big).$$

(b) 存在一个通用常数 $C_1 > 0$，使得对于任意 $t \geq 0$,

$$\mathbb{E}\big[\varepsilon^t\big] = 0, \quad \mathbb{E}\big[\big|\varepsilon^t\big|^2\big] = \sigma_d^2 \quad and \quad \mathbb{E}\big[\big|\varepsilon^t\big|^4\big] = C_1\sigma_d^4.$$

噪声方差 σ_d^2 在我们的尺度极限中起着关键作用。特别是，它允许随 d 增长，其增长速率决定了随机梯度下降迭代的精确标度极限。为方便起见，下面省略下标 d。现在我们介绍关于协方差光滑性以及随机过程 W 的高阶矩的主要假设。

假设 1.2 (光滑性条件**).** 随机过程 $\{W(s,x)\}_{s\geq 0, x\in[0,1]}$ 满足以下光滑性条件:

1. 函数 $A\colon [0,1]^2 \to \mathbb{R}$ 满足以下条件：对于任意 $x \in [0,1]$, $A(x,\cdot) \in C^1([0,1])$。因此，存在常数 C_2, C_3，使得对于任意 $x, y, z \in [0,1]$，有

$$|A(x,y)| \leq C_2 \qquad and \qquad |A(x,z) - A(y,z)| \leq C_3|x-y|.$$

我们进一步假设存在一个常数 C_4，使得

$$|A(x,x) + A(y,y) - 2A(x,y)| \leq C_4^2|x-y|^2.$$

2. 对于函数 B： $[0,1]^4 \to \mathbb{R}$，存在常数 C_5、C_6、$C_7 > 0$，使得对于任意x_1、x_2、x_3、$x_4 \in [0,1]$、$|B(x_1, x_2, x_3, x_4)| \le C_5$ 以及

$$\begin{aligned}&|B(x_1,x_3,x_1,x_3)+B(x_2,x_3,x_2,x_3)-2B(x_1,x_3,x_2,x_3)| \le C_6^2|x_1-x_2|^2,\\&\big|B(x_1,x_1,x_1,x_1)+B(x_2,x_2,x_2,x_2)+6B(x_1,x_1,x_2,x_2)-4B(x_1,x_1,x_1,x_2)\\&\qquad -4B(x_1,x_2,x_2,x_2)\big| \le C_7^4\big|x_1-x_2\big|^4.\end{aligned}$$

3. 对于函数 E： $[0,1]^8 \to \mathbb{R}$，存在常数 C_8, C_9，使得对于所有 $x_1, \cdots, x_8 \in [0,1]$，$E(x_1, \cdots, x_8) \le C_8$ 以及

$$\begin{aligned}&\big|E(x_1,x_1,x_1,x_1,x_3,\cdots,x_6)+E(x_2,x_2,x_2,x_2,x_3,\cdots,x_6)\\&+6E(x_1,x_1,x_2,x_2,x_3,\cdots,x_6)-4E(x_1,x_1,x_1,x_2,x_3,\cdots,x_6)\\&\qquad -4E(x_1,x_2,x_2,x_2,x_3,\cdots,x_6)\big| \le C_9^4\big|x_1-x_2\big|^4.\end{aligned}$$

上述假设的动机。假设 1.1 和 1.2旨在确保（3）中的插值过程具有紧性。特别是，关于二阶和四阶矩的假设是推导建立尺度极限所需紧性条件的必要条件，其中二阶矩信息实际上会出现在极限中。此外，关于八阶矩的假设是推导建立涨落所需紧性条件的必要条件。在这种情况下，二阶矩信息出现在漂移项中，而二阶和四阶矩信息都会出现在扩散项中。

从统计学角度看，我们上述的假设允许数据和噪声序列服从相对广泛的一类分布。具体而言，我们现在刻画 A 的特征值衰减与假设 1.2 之间的关系。我们通过将中心化随机过程 $\{W(s, x)\}_{s\ge 0,\ x\in[0,1]}$ 的协方差函数 A 写成如下 Mercer 分解形式来进行分析，

$$A(x,y)=\sum_{k=1}^{\infty}\lambda_k\phi_k(x)\phi_k(y), \tag{4}$$

其中 $\{\lambda_k\}_{k\ge 1}$ 是按降序排列的正系数序列（特征值），而 $\{\phi_k\}_{k\ge 1}$是 $L^2([0,1]) \cap C^1([0,1])$ 的标准正交基。在下面的结果中，我们给出关于特征值的一个充分条件，使得假设 1.2-1 得以满足。

命题 1.3。 假设 $\{\phi_k\}_{k\ge 1} \subset L^2([0,1]) \cap C^2([0,1])$ 是 $L^2([0,1])$的标准正交基。定义

$$S_{N,0} := \sum_{k=1}^{N}\lambda_k\phi_k(x)\phi_k(y) \quad and \quad S_{N,1} := \sum_{k=1}^{N}\lambda_k\phi_k(x)\phi_k'(y),$$

其中 ϕ_k' 表示 是 ϕ_k 的导数。如果 $\{(\lambda_k, \phi_k)\}_{k\ge 1}$满足对于所有 $k \ge 1$和 $i = 0, 1$，

$$\sup_{x\in[0,1]}|\phi_k^{(i)}(x)| \le C_{k,i}, \qquad \sum_{k=1}^{\infty}\lambda_k C_{k,i}^2 < \infty, \qquad \sum_{k=1}^{\infty}\lambda_k C_{k,0}C_{k,2} < \infty \tag{5}$$

则 $\{S_{N,0}\}_{N\ge 1}$ 一致有界，且 $\{S_{N,1}\}_{N\ge 1}$ 一致有界且一致等度连续。因此，随机过程 $\{W(s, x)\}_{s\ge 0,\ x\in[0,1]}$，其协方差函数为 A （见(4)），满足假设1.2-1。

命题1.3的证明。由（ $\{S_{N,0}\}_{N\geq 0}$ 和 $\{S_{N,1}\}_{N\geq 0}$ 一致有界这一事实可直接由（5）得出。现在，对任意 x, y_1, $y_2 \in [0,1]$,

$$|S_{N,1}(x,y_1)-S_{N,1}(x,y_2)| \leq \sum_{k=1}^{N} \lambda_k C_{k,0} |\phi_k'(y_1)-\phi_k'(y_2)| \leq \big(\sum_{k=1}^{\infty} \lambda_k C_{k,0} C_{k,2}\big)|y_1-y_2|.$$

因此， $\{S_{N,1}\}_{N\geq 0}$ 是一致等度连续的。因此，根据阿尔泽拉-阿斯科利定理， $A(x,\ \cdot) \in C^1([0,1])$。

由于对称性， A 有界，且关于每个变量都是利普希茨连续的，并且

$$\begin{aligned}|A(x,x)+A(y,y)-2A(x,y)| &= \lim_{N\to\infty} |S_{N,0}(x,x)+S_{N,0}(y,y)-2S_{N,0}(x,y)| \\ \leq \lim_{N\to\infty} \sum_{k=1}^{N} \lambda_k \big(\phi_k(x)\phi_k(x)+\phi_k(y)\phi_k(y)-2\phi_k(x)\phi_k(y)\big) &\leq \big(\sum_{k=1}^{\infty} \lambda_k C_{N,1}^2\big)|x-y|^2.\end{aligned}$$

因此，假设 1.2-1 成立。 ■

备注 1.4。 我们在此强调，对于命题1.3，我们既不作高斯性假设。虽然我们主要讨论了假设 1.2-1，但若进一步假设中心化随机过程 $\{W(s,\ x)\}_{s\geq 0,\ s\in[0,1]}$ 是椭圆过程，则根据广义伊瑟利斯定理可直接得出，假设 1.2-2 和假设 1.2-3 也成立。或者，也可以像上文那样进行更明确的计算，以关联假设 1.2-2 和假设 1.2-3 中的特征值与光滑性条件。总之，我们的假设基于样本分布的 k^{th} 阶矩的空间结构，对于 $k=2$，4，8，只要协方差的特征值衰减不太慢，这些假设就成立。

备注 1.5。 特征值衰减假设通常用于分析固定有限维和固定无限维设定下的最小二乘随机梯度下降算法；例如，参见 Lin 和 Rosasco (2017年); Dieuleveut 等人 (2017年);Pillaud-Vivien 等人 (2018年); M¨ucke 等人 (2019年);Bartlett 等人 (2020年)。事实上，这些假设最初是在分析核岭回归估计量（即再生核希尔伯特空间中的无限维最小二乘）时提出的，见 Caponnetto 和 De Vito (2007)。从实际统计数据分析的角度来看，这些假设的动机来自经验观察：大多数大型数据集都呈现特征值衰减现象，并且可以用低秩矩阵很好地近似（Udell 和 Townsend，2019年）。

特别地，(4) 和 (5) 可用于刻画涵盖一类实践中广泛使用的数据集的实例。假设数据协方差矩阵 Σ_d 具有如下特征值分解：

$$\Sigma_d = \sum_{k=1}^{\infty} \lambda_k \mathbf{v}_k \mathbf{v}_k^{\intercal}, \tag{6}$$

其中 $\{\mathbf{v}_k\}_{k\geq 1}$ 是 $\mathbb{R}^d$ 的标准正交基，满足 $\mathbf{v}_k^{\intercal} = (\phi_k(i/d))^{i\in[d]}$ 对所有 $k \geq 1$均成立。于是，当特征值序列 $\{\lambda_k\}_{k\geq 1}$ 衰减得不太慢时，假设（5）成立。例如，当 $\{\phi_k\}_{k\geq 1}$ 是 $L^2([0,1])$ 的傅里叶基时，只要 5）中的条件成立，那么当 $\{\lambda_k\}_{k\geq 1}$ 以 $k^{-5/2-\varepsilon}$ 的速度衰减时，这些条件对任意小的正数 ε均成立。为简便起见，高斯情形下的具体计算见第 2.2 节和引理2.4 。

我们还强调，我们的分析具有足够的普适性，允许数据序列 $\{x^t\}_{t\geq 0}$ 和误差序列 $\{\varepsilon^t\}_{t\geq 0}$ 内部存在一定程度的依赖性。此类假设

通常用于若干应用领域，包括强化学习（Meyn，2022年）以及序贯或在线决策（Chen 等，2021年；Khamaru 等，2021年）。此外，我们也可以允许数据序列与误差序列之间存在一定程度的依赖性。为使论述更简洁，我们不详细讨论这些扩展。

1.3 主要结果

本节介绍最小二乘在线随机梯度下降轨迹及其涨落的尺度极限。主要结果中的极限满足随机微分方程 [1]，这些方程由不同的高斯随机场过程驱动。按照惯例，对于任意高斯随机场过程 $\{\xi(s,\ x)\}_{s\geq 0,\ x\in[0,1]}$，$\mathrm{d}\xi(s,\ x)$ 表示该高斯随机场过程关于时间变量 s 的无穷小增量。

关于极限随机过程的初始条件，我们还采用以下约定：这些初始条件来自第 1 节中引入的时空插值函数的极限。在 $t=0$ 处，我们仅在空间上进行插值，并将分段线性插值函数记为 Θ^d。在以下各项结果中关于 $\{\Delta\theta_i^0\}_{i\in[d]}$ 的假设下，Θ^d 的子序列极限存在于$C([0,1])$ 中；我们选取其中一个子序列极限作为初始条件，记作 Θ_0。

1.3.1 尺度极限

现在，我们陈述不同噪声方差阶下最小二乘在线随机梯度下降的尺度极限结果。

定理 1.6（尺度极限）。设初始条件满足：对于任意 $0\leq i\leq d$，存在常数 R, L，使得

$$\mathbb{E}\left[|\Delta\theta_i^0|^4\right]\leq R^4 \qquad and \qquad \mathbb{E}\left[|\Delta\theta_i^0-\Delta\theta_{i-1}^0|^4\right]\leq L^4 d^{-4}.$$

此外，设假设 *1.1* 和假设 *1.2* 成立，并令 $\{\xi_1(s,x)\}_{s\in[0,\tau],\ x\in[0,1]}$ 表示一个中心化高斯随机场过程，使得对于任意 $s_1, s_2\in[0,\tau]$ 和 $x, y\in[0,1]$，

$$\mathbb{E}[\xi_1(s_1,x)\xi_1(s_2,y)]=\sigma_1(s_1,s_2,x,y) \quad with \quad \sigma_1(s_1,s_2,x,y)=(s_1\wedge s_2)A(x,y). \tag{7}$$

进一步，设存在一致常数 $C_{s,\ 1}$，使得对于极限序列中的所有 $d\geq 1, T>0$，$\max(\eta dT, \sigma\eta T^{\frac{1}{2}})\leq C_{s,\ 1}$。那么，我们得到以下尺度极限。

1. 低噪声，即 $\sigma^2/d^2T\to 0$：进一步假设存在一致正常数α，使得 $\lim_{d,\ T\to\infty}\eta dT=\alpha$。那么，对于任意 $\tau\in(0,\infty)$，$\{\Theta^{d,T}\}_{d\geq 1,\ T>0}$弱收敛到函数 Θ，当 $d, T\to\infty$ 时成立。此外，极限 $\Theta\in C^1([0,\tau];\ C^1([0,1]))$ 是以下常微分方程的唯一连续解：

$$\partial_s\Theta(s,x)=-\alpha\int_0^1 A(x,y)\Theta(s,y)\mathrm{d}y, \quad s\in[0,\tau], x\in[0,1]. \tag{8}$$

2. 中等噪声，即 $\sigma^2/d^2T\to(0,\infty)$：假设还存在一致正常数 α,β，使得 $\lim_{d,T\to\infty}\sigma^2/(d^2T)=\beta^2$ 且 $\lim_{d,T\to\infty}\eta dT=\alpha$（因此，作为推论 $\lim_{d,T\to\infty}\eta\sigma T^{\frac{1}{2}}=\alpha\beta$）。那么，对任意 $\tau\in(0,\infty)$，$\{\Theta^{d,T}\}_{d\geq 1,T>0}$ 弱收敛

[1]在某些情况下，随机微分方程极限会退化为常微分方程极限。

到随机过程 $\{\Theta(s,\cdot)\}_{s\in[0,\tau]}$，当 $d,T\to\infty$ 时。此外，极限 Θ 是 $C([0,\tau];C([0,1]))$ 中的唯一解，满足以下随机微分方程：

$$\mathrm{d}\Theta(s,x)=-\alpha\int_0^1 A(x,y)\Theta(s,y)\mathrm{d}y\mathrm{d}s+\alpha\beta\mathrm{d}\xi_1(s,x),\quad s\in[0,\tau],x\in[0,1]. \tag{9}$$

3. 高噪声，即 $\sigma^2/d^2T\to\infty$：假设还存在一个一致正常数α，使得 $\lim_{d,\ T\to\infty}\eta\sigma T^{\frac{1}{2}}=\alpha$。那么对于任意 $\tau\in(0,\ \infty)$，$\{\Theta^{d,T}\}_{d\geq1,\ T>0}$弱收敛到随机过程 $\{\overline{\Theta}(s,\ \cdot)\}_{s\in[0,\tau]}$，当 $d,\ T\to\infty$。此外，极限 Θ 是以下随机微分方程在$C([0,\tau];\ C([0,1]))$ 中的唯一解：

$$\mathrm{d}\Theta(s,x)=\alpha\mathrm{d}\xi_1(s,x),\quad s\in[0,\tau],x\in[0,1]. \tag{10}$$

注 1.7。 定理 1.6 表明，在高维缩放下，随机梯度下降轨迹的极限行为呈现三阶段相变：在低噪声情形下，它表现为弹道态（即由无限维常微分方程刻画）；在中等噪声情形下，它转为扩散型（即由无限维随机微分方程刻画）；而在高噪声情形下，则变为纯随机。该三阶段相变的边界得到了精确刻画，并明确了其对维度阶数、迭代次数和步长选择的依赖关系。在中等噪声和高噪声情形下，扩散项 ξ_1 的协方差由假设 A 1.2中的二阶矩函数 A 确定。我们还指出，初始条件是按坐标给出的，且相当宽松。

备注 1.8。在低噪声情形下，根据 (8)，我们看到，对于任意 $s\in[0,\tau]$ 和 $y\in[0,1]$，$\Theta(s,\cdot)$ 与 $A(\cdot,y)$ 具有相同的光滑阶数。这一现象更具一般性，即如果进一步假设，对于某个有限正整数 k 以及任意 $y\in[0,1]$，$A(\cdot,y)\in C^k([0,1])$，那么可以证明 $\Theta\in C^1([0,\tau];C^k([0,1]))$。

备注 1.9。 假设 1.2 是命题 4.1和命题4.2中证明紧性结果所必需的，因此也是证明弱极限存在性的必要条件。值得一提的是，在低噪声设定下，当噪声方差满足 $\sigma^2=O(d^2)$ 时，假设 1.2 可以通过删除关于 B 的最后一个条件以及关于 E的条件来放宽。在放宽后的假设下，可以用相同的方法证明紧性结果，只需在命题 4.1 和命题4.2的证明中考虑二阶矩界，而非四阶矩界。为简化分析并保持一致性，定理 1.6 的证明基于假设 1.2。

1.3.2 涨落

现在我们研究低噪声情形下 $\{\Delta\theta_i^t\}_{i\in[d],0\leq t\leq N}$ 的涨落。为此，我们考察 $\{\Delta\theta_i^t\}_{i\in[d],0\leq t\leq N}$ 与其标度极限之间的一个重新缩放后的差值。具体而言，对于任意 $s\in[0,\tau]$、$x\in[0,1]$，定义

$$U^{d,T}(s,x):=\gamma\big(\overline{\Theta}^{d,T}(s,x)-\Theta(s,x)\big), \tag{11}$$

其中，γ 是缩放参数；我们预期当 $d,T\to\infty$ 时，$\gamma\to\infty$。现在我们陈述涨落结果。

定理 1.10（涨落）。 设初始条件满足：对于任意 $0\leq i\leq d$，存在一致常数 D、M，使得

$$\mathbb{E}\left[|U^{d,T}(0,\frac{i}{d})|^4\right]\leq D^4\qquad and\qquad \mathbb{E}\left[|U^{d,T}(0,\frac{i}{d})-U^{d,T}(0,\frac{i-1}{d})|^4\right]\leq M^4d^{-4}.$$

令 $\{\xi_2(s、x)\}_{s\in[0,\tau],\ x\in[0,1]}$ 和 $\{\xi_3(s、x)\}_{s\in[0,\tau],\ x\in[0,1]}$ 表示两个相互独立的中心化高斯场过程，使得对于任意 $s_1、s_2\in[0,\tau]$ 和 $x、y\in[0,1]$,

$$\mathbb{E}[\xi_2(s_1,x)\xi_2(s_2,y)]=\sigma_2(s_1,s_2,x,y)\quad with\quad \sigma_2(s_1,s_2,x,y)=(s_1\wedge s_2)A(x,y), \tag{12}$$

$$\mathbb{E}[\xi_3(s_1,x)\xi_3(s_2,y)]=\sigma_3(s_1,s_2,x,y)\quad with$$
$$\sigma_3(s_1,s_2,x,y)=\int_0^{s_1\wedge s_2}\int_0^1\int_0^1\Theta(s,z_1)\Theta(s,z_2)\tilde{B}(x,z_1,x,z_2)\mathrm{d}z_1\mathrm{d}z_2\mathrm{d}s, \tag{13}$$

其中，对于任意 $x_1,x_2,x_3,x_4\in[0,1]$，$B(x_1,x_2,x_3,x_4)=B(x_1,x_2,x_3,x_4)-A(x_1,x_3)A(x_2,x_4)$。此外，假设低噪声设定下定理 *1.6* 中的条件，即 $\sigma^2/d^2T\to0$，均成立。令 $\eta=\frac{\alpha}{dT}$，并假设存在一致常数 $C_{s,2}$，使得沿极限序列对所有 $d\geq1,T>0$，均有 $max(\gamma T^{-\frac{1}{2}},\gamma d^{-1},\gamma\sigma d^{-1}T^{-\frac{1}{2}})\leq C_{s,2}$。于是得到以下涨落结果。

1. ***P***文章间的相互作用占主导：进一步假设，当 $d,T\to\infty$ 时，有

$$T=o(d^2),\quad \sigma=O(d),\quad \gamma T^{-\frac{1}{2}}\to\zeta\in(0,\infty),\quad and\quad \gamma\sigma d^{-1}T^{-\frac{1}{2}}\to\beta\in[0,\infty),$$

对于某些一致常数 ζ,β，上述结论成立。于是，对任意 $\tau\in(0,\infty)$，随机梯度下降粒子 $\{U^{d,T}\}_{d\geq1,T>0}$ 的涨落依分布收敛到函数 $U\in C([0,\tau];C([0,1]))$，当 $d,T\to\infty$ 时。此外，对任意 $s\in[0,\tau],x\in[0,1]$，极限 U 是随机微分方程在 $C([0,\tau];C([0,1]))$ 中的唯一解

$$\begin{aligned}\mathrm{d}U(s,x)=&-\alpha\int_0^1A(x,y)U(s,y)\mathrm{d}y\mathrm{d}s+\alpha\beta\mathrm{d}\xi_2(s,x)\\&+\alpha\zeta\mathrm{d}\xi_3(s,x),\quad s\in[0,\tau],x\in[0,1].\end{aligned} \tag{14}$$

2. 噪声占主导：进一步假设，当 $d,T\to\infty$ 时，我们有

$$\max(d,T^{\frac{1}{2}})\ll\sigma\ll dT^{\frac{1}{2}},\qquad and\qquad \gamma\sigma d^{-1}T^{-\frac{1}{2}}\to\beta\in(0,\infty),$$

对于一致常数 β。那么，对任意 $\tau\in(0,\infty)$，随机梯度下降粒子的涨落 $\{U^{d,T}\}_{d\geq1,T>0}$ 依分布收敛到函数 $U\in C([0,\tau];C([0,1]))$，当 $d,T\to\infty$ 时。此外，对任意 $s\in[0,\tau],x\in[0,1]$，极限 U 是 $C([0,\tau];C([0,1]))$ 中的唯一解，该解满足随机微分方程

$$\mathrm{d}U(s,x)=-\alpha\int_0^1A(x,y)U(s,y)\mathrm{d}y\mathrm{d}s+\alpha\beta\mathrm{d}\xi_2(s,x),\quad s\in[0,\tau],x\in[0,1]. \tag{15}$$

注 1.11. 定理中的设定1.10进一步将定理 1.10中的低噪声区间划分为两个子区间。前两个区间对应于导致涨落的主导项分别由粒子相互作用（其期望由协方差函数 A刻画）和噪声方差所致的情形。特别地，当粒子相互作用占主导时，会出现两个相互独立的扩散项；而过程ξ_2 的协方差仅由假设 A 中的二阶矩函数 1.2决定，第二个扩散项中的过程 ξ_3 的协方差则由 A 和四阶矩函数 B共同决定。完整的极限识别结果见定理 3.5。

注 1.12. 除了定理中引入的两个子区间1.10之外，还存在另一个子区间，其中 **插值误差占主导**。具体而言，在定理 1.10的假设下，若进一步假设当 d 时， $T \to \infty$，则有

$$d = O(T^{\frac{1}{2}}), \qquad \sigma = O(T^{\frac{1}{2}}) \qquad \text{and} \qquad \gamma d^{-1} \to 0,$$

那么，对于任意 $\tau \in (0, \infty)$，随机梯度下降粒子 $\{U^{d,T}\}_{d \geq 1, T>0}$的涨落依分布收敛到函数 $U \in C([0,\tau]; C([0,1]))$，当 $d, T \to \infty$ 时。此外，对于任意 $s \in [0,\tau]$, $x \in [0,1]$，极限 U 是以下常微分方程（具有随机初始条件）的唯一连续解

$$\mathrm{d}U(s,x) = -\alpha \int_0^1 A(x,y)U(s,y)\mathrm{d}y\mathrm{d}s, , \quad s \in [0,\tau], x \in [0,1]. \tag{16}$$

这个子区域源于需要用黎曼和近似右侧积分（8），因此该区域对应于插值误差占主导的情形。在此区域中，我们选取较小阶的 γ 以确保插值误差消失。这样做会使粒子相互作用和噪声引起的涨落在极限中受到抑制。因此，对于涨落，我们得到如（16）所示的常微分方程的退化收敛。

然而，由于 γ 的阶较小，且满足定理1.10中的假设，式（16）的初始化是平凡的，因此式（16）的解也同样平凡。这表明，选择 γ 阶为o（d）的量过小，以至于抑制了极限中的所有涨落。值得指出的是，值得关注的情形是 $\gamma = \Theta$（d），此时插值误差不会消失。然而，要确定这种情形下的极限，就需要确定黎曼和的数值积分误差的极限，这超出了本文的研究范围。

注 1.13。我们强调，对于定理1.6、定理1.10 以及注1.12中的每个方程，都存在一个使该方程成立的概率空间。此类概率空间的存在性可由斯科罗霍德表示定理推出。更多细节请参阅第3节。

注 1.14. 定理 1.6和 1.10 中分别建立的尺度极限与涨落，经过少量调整后，同样适用于 Fang 等人, 2018年，方程（7）和（8））中提出的随机梯度下降在线乘子自助法版本；这些结果可用于实际的高维统计推断。

1.4 相关工作

现在，我们将研究结果置于更广泛的随机梯度下降分析文献背景中 s.

随机梯度下降的高维尺度极限。 Ben Arous 等人 (2022年) 在统计模型框架下，研究了一类非凸问题中高维情形下在线随机梯度下降的尺度极限。他们推导了在线随机梯度下降有限维汇总统计量的尺度极限，但未给出任何涨落结果。此外，维度与迭代次数之间的精确尺度关系，以及数据生成过程（尤其是协方差结构）的影响，尚未得到探讨。 Wang 等人 (2017年) 分析了用于最小二乘回归和主成分分析的在线随机梯度下降算法，并推导出经验密度作为偏微分方程解的尺度极限。然而，他们的分析仅限于各向同性协方差矩阵这一特殊情形，且未提供任何涨落结果。另请参见 Wang 和 Lu (2019年); Veiga 等人 (2022年)，了解特定模型的相关研究。高维

最小二乘问题中多轮（及在线）随机梯度下降的动力学由Paquetteand Paquette (2021年); Paquette 等人 (2022a,b) 进行了研究。在不假设存在真实底层统计模型的情况下，研究表明，在对输入数据作出相当一般的（非高斯）假设时，某个齐次化随机微分方程能够精确刻画均方损失。然而，这些研究并未识别出任何相变现象（从弹道态到扩散行为），仅适用于随机梯度下降迭代的一种特定二次统计量，也未刻画其涨落。

关于Celentano 等人（2021年）和Gerbelot 等人（2022年）的研究也刻画了高维情形下某类优化问题中随机梯度下降变体的渐近行为。他们的方法基于统计物理中所谓的动力学平均场理论。不同于研究期望目标函数上的在线随机梯度下降的本文，Celentano 等人（2021年）和 Gerbelot 等人（2022年）研究的是有限和目标函数（或经验风险最小化）上的小批量随机梯度下降。Celentano 等人（2021年）的分析要求输入为各向同性次高斯输入。尽管Gerbelot 等人（2022年）允许非各向同性协方差，但其要求输入满足高斯性假设。最后，随着维度增长，他们仅追踪轨迹的一个实值泛函，且未提供涨落结果。

据我们所知，我们的工作首次刻画了特定问题——具有光滑协方差结构的最小二乘回归——中在线随机梯度下降的完整无限维轨迹极限及相关涨落。

SGD 的其他高维分析。随机矩阵理论也用于分析基于梯度的迭代算法（Deift and Trogdon, 2019年；Ding and Trogdon, 2022年）。在机器学习领域，Paquette et al. (2021, 2022c) 分析了用于高维最小二乘线性回归的全批量和小批量 SGD，并将其极限确立为沃尔泰拉方程的解。Ba et al. (2022) 研究了利用随机矩阵理论分析神经网络全批量梯度下降动力学的结果。Chandrasekher et al. (2022, 2023) 使用高斯过程技术研究了适用于某些高维非凸问题的小批量 SGD。他们的研究高度依赖各向同性和高斯性假设。Tan and Vershynin (2023) 在各向同性高斯性假设下，采用状态空间方法分析了在线 SGD 的高维情形。Ben Arous et al. (2021) 的研究也采用了类似方法，在信号恢复设定中建立了高概率界。最近，Agrawalla et al. (2023) 建立了在线最小二乘 SGD 的高维正态近似结果，而 Durmus et al. (2021, 2022) 则建立了相应的尾界。

近年来，关于过参数化神经网络随机梯度下降的平均场分析也得到了深入研究。然而，这类工作在假设参数维度增长的同时，却假设数据维度固定。鉴于近期这方面涌现了大量研究，无法在此一一列举。关于相关研究，可参阅 Chizat 和 Bach (2018年);Mei 等人(2018年); Sirignano 和 Spiliopoulos(2020a,b);Pham 和 Nguyen (2021年); Rotskoff 和 Vanden-Eijnden (2022年); Ding 等人(2022年);Sirignano 和 Spiliopoulos (2022年); Abbe 等人 (2022年); Gess 等人 (2022年)，可从中选取部分文献作为参考。

固定维度随机梯度下降分析。在固定维度设定下研究随机梯度下降的扩散极限，是一个经典课题。关于这一主题的教科书式论述，可参见Kushner 和 George Yin (2003年); Kushner 和 Clark (2012年); Benveniste 等人 (2012年); Ljung 等人 (2012年。这类研究的核心思想是，在特定的缩放下，证明经过适当时间插值的随机梯度下降迭代收敛到一个多维奥恩斯坦-乌伦贝克过程。最近，Li 等人 (2019年提出了一个框架，用随机微分方程来近似一类相对一般的随机梯度算法的动态。另请参见Krichene 和 Bartlett

(2017年); Li 等人 (2018年); Gupta 等人 (2020年); Fontaine 等人 (2021年)，其他近期相关研究的部分列表可参见这些文献。

固定维度情形下随机梯度下降的几乎必然收敛和中心极限定理也已得到广泛研究。关于几乎必然收敛结果，参见 Mertikopoulos 等人（2020年）、Sebbouh 等人（2021年）、Liu 和 Yuan（2022年）及其中的参考文献。关于中心极限定理，相关工作的部分列表包括 Polyak 和 Juditsky（1992年）、Ruppert（1988年）、Duchi 和 Ruan（2020年）、Toulis 和 Airoldi（2017年）、Asi 和 Duchi（2019年）、Yu 等人（2021年）、Dieuleveut 等人（2020年）、Barakat 等人（2021年）、Davis 等人（2023年）及其中的参考文献。我们还特别指出 Anastasiou 等人（2019年）和 Shao 和 Zhang（2022年）的研究，他们建立了随机梯度下降的非渐近正态近似。关于随机梯度下降及其变体的期望界和高概率界综述，参见 Bottou 等人（2018年）和 Lan（2020年）。

其他算法的尺度极限。 我们指出，随机游走 Metropolis（RWM）算法、未校正朗之万算法（ULA）和 Metropolis 校正朗之万算法（MALA）等迭代采样算法的高维尺度极限已得到充分研究。例如，Pillai 等人(2012年) 和 Mattingly 等人(2012年) 将尺度极限刻画为无限维随机微分方程（或等价地，随机偏微分方程）；相关方向的其他研究可参见其中的参考文献。尽管这些研究在理念上与我们的方法相关，但其中考察的采样算法对应的设定与我们的不同，因为其相互作用并非由任何数据生成过程刻画。近年来，聚焦于高维线性回归，Qiu 和 Sen （2022年）和 Mukherjee 和 Sen （2022年）研究了变分推断算法的尺度极限，而这些算法与随机梯度下降并无直接关联。

补充讨论。如前所述，本文研究的是长程相互作用区间。另一种情形是协方差矩阵较为“粗糙”，其各元素之间可能发生一定程度的突变（例如，允许 Σ_d 中出现各种结构化稀疏模式），这对应于短程相互作用区间。过去几十年来，大量关于相互作用粒子系统的研究聚焦于相互作用粒子系统中涨落的尺度极限与 Kardar-Parisi-Zhang（KPZ）方程之间的联系。例如，Bertini 和 Giacomin（1997）；Bal´azs 和 Sepp¨al¨ainen（2010）；Quastel（2011）；Corwin（2012）；Corwin 等人（2018）；Ghosal（2017）；Corwin 等人（2020）。其中许多研究证明，在弱噪声缩放下，相关粒子系统的高度函数涨落会收敛到 KPZ 方程。少数研究（包括一些近期的突破性成果）（Matetski 等人，2021；Dauvergne 等人，2022；Quastel 和 Sarkar，2023）表明，在通常所说的 KPZ 空间、时间和涨落缩放下，系统会收敛到 KPZ 固定点。本文关于长程相互作用情形的涨落结果，在极限下不会得到 KPZ 方程或 KPZ 固定点。然而，我们确实认为，即使对于在线随机梯度下降求解最小二乘这一看似简单的问题，在短程区间内，尺度极限也会导向这些前所未见的极限行为。在一篇即将发表的工作中，我们将详细研究短程相互作用情形。

2 应用与示例

作为定理 1.6 和 1.10中主要结果的一项应用，本节将展示如何利用这些结果研究最小二乘在线随机梯度下降的某些具体性质，例如均方误差（MSE）和预测误差（PE）。

均方误差（MSE）和预测误差（PE）： 对于给定的 d 和 T，最小二乘在线随机梯度下降估计量在

时间 $s \in [0,\tau]$ 处的时间插值均方误差和预测误差分别定义为

$$\mathsf{MSE}^{d,T}(s) := \frac{1}{d}\sum_{i=1}^{d} \overline{\Theta}^{d,T}(\frac{\lfloor sT \rfloor}{T}, \frac{i}{d})^2,$$

$$\mathsf{PE}^{d,T}(s) := \frac{1}{d^2}\sum_{i,j=1}^{d} A(\frac{i}{d},\frac{j}{d})\overline{\Theta}^{d,T}(\frac{\lfloor sT \rfloor}{T}, \frac{i}{d})\overline{\Theta}^{d,T}(\frac{\lfloor sT \rfloor}{T}, \frac{j}{d}).$$

具体而言，我们首先根据方程（8）、（9）和（10）的解，计算均方误差 和 预测误差 的高维尺度极限，并展示低噪声设定下极限均方误差 和极限预测误差 的衰减性质。接着，我们根据方程（8）以及方程（14）、（15）和（16）的解，计算均方误差 和预测误差 的涨落。

2.1 均方误差 与 预测误差的尺度极限与涨落

利用定理 1.6，我们得到最小二乘在线随机梯度下降的均方误差 和预测误差 的尺度极限如下。具体而言，其呈现出依赖于噪声水平的三阶段相变。接下来，我们将极限均方误差 和 预测误差 定义如下：

$$\mathsf{MSE}(s) := \int_0^1 \Theta(s,x)^2 dx, \qquad \mathsf{PE}(s) := \int_0^1\int_0^1 \Theta(s,x)A(x,y)\Theta(s,y)dxdy. \tag{17}$$

命题 2.1。 假设定理*1.6* 中的假设成立。则有以下收敛：*(i)* 均方误差$^{d,\ T}(s)$ $\to$均方误差（s）以及 *(ii)* 预测误差$^{d,\ T}(s)$ $\to$预测误差（s）（在低噪声情形下依概率收敛，在另外两种情形下依分布收敛），前提是满足以下缩放条件之一：

(1) 低噪声： $\sigma d^{-1}T^{-\frac{1}{2}} \to 0,\ \eta dT \to \alpha \in (0,\infty)$ 当 $d,T \to \infty$ 且 Θ 是下式的解 (8)。

(2) 中等噪声： $\sigma d^{-1}T^{-\frac{1}{2}} \to \beta \in (0,\infty),\ \eta dT \to \alpha \in (0,\infty)$ 当 $d,T \to \infty$ 且 Θ 是下式的解 (9)。

(3) 高噪声： $\sigma d^{-1}T^{-\frac{1}{2}} \to \infty,\ \eta\sigma T^{\frac{1}{2}} \to \alpha \in (0,\infty)$ 当 $d,T \to \infty$ 且 Θ 是下式的解(10)。

接下来，我们分别刻画均方误差^{d,T}和预测误差^{d,T}的涨落，即 γ (均方误差^{d,T} − 均方误差) 和 γ (预测误差^{d,T} − 预测误差)。

命题 2。 **2.** 假设定理 *1.6* 和定理*1.10* 中的假设成立。那么

(i) γ (均方误差$^{d,\ T}(s)$ − 均方误差(s)) $\to 2\int_0^1 \Theta(s,x)U(s,x)\mathrm{d}x$ 依分布收敛，

(ii) $\gamma\left(\mathsf{PE}^{d,T}(s) - \mathsf{PE}(s)\right) \to 2\int_0^1\int_0^1 \Theta(s,x)A(x,y)U(s,y)\mathrm{d}x\mathrm{d}y$ *in distribution,*

满足以下任一缩放条件：

(1) $\gamma = \zeta T^{-\frac{1}{2}}$ 对于某个 $\zeta \in (0,\infty)$. $T = o(d^2)$ $\sigma = O(d)$ 当 d时， $T \to \infty$ 且 Θ， U分别是 (8) 和 (14) 的解。

(2) $\gamma = \beta\sigma^{-1}dT^{-\frac{1}{2}}$ 对于某个 $\beta \in (0,\infty)$. $\max(d, T^{\frac{1}{2}}) \ll \sigma \ll dT^{\frac{1}{2}}$ 当 d时， $T \to \infty$ 且 Θ， U分别是 (8) 和 (15)的解。

(3) $1 \ll \gamma \ll d$， $d = O(T^{\frac{1}{2}})$， $\sigma = O(T^{\frac{1}{2}})$ 在 d、$T \to \infty$和 Θ中， U分别是 (8) 以及相应 (16) 的解。

2.2 补充见解与具体示例

本节中，我们将证明，在假设 1.1 （a）部分）和假设 1.2下，定义 A、B、E 中的函数 1.2 是有限维数据相应矩的极限描述。我们首先证明，由有限维协方差矩阵 Σ_d 及数据 $\{x_i^t\}_{1\le i\le d}$ 所诱导的分段常值函数 A 在定义 1.2中定义的函数上一致收敛。在陈述收敛结果之前，我们先引入一些必要的定义和记号。

定义 2.3。给定一个实数元素对称矩阵 $d\times d$，定义一个分段常值函数于 $[0,1]^2$ 上：将 $[0,1]^2$ 划分为 d^2 个边长为 $1/d$ 的小正方形，并令

$$W_\Sigma(x,y) \coloneqq \Sigma(i,j) \text{ if } \lceil dx\rceil = i, \lceil py\rceil = j.$$

回顾一下，Σ_d 表示数据 $\{x_i^t\}_{1\le i\le d}$的协方差矩阵。容易看出，当 $d\to\infty$时，在假设1.1 和假设1.2下，我们有

$$\|W_{\Sigma_d} - A\|_\infty \coloneqq \sup_{x,y\in[0,1]} |W_{\Sigma_d}(x,y) - A(x,y)| \to 0.$$

事实上，由定义2.3可得

$$W_{\Sigma_d}(x,y) - A(x,y) = \begin{cases} 0, & \lceil dx\rceil, \lceil dy\rceil \in \{0,1,\cdots,d\} \\ A\Big(\frac{\lceil dx\rceil}{d}, \frac{\lceil dy\rceil}{d}\Big) - A(x,y), & \text{otherwise.} \end{cases}$$

由假设 1.2可得

$$\begin{aligned} &\Big|A\Big(\frac{\lceil dx\rceil}{d}, \frac{\lceil dy\rceil}{d}\Big) - A(x,y)\Big| \\ \le &\Big|A\Big(\frac{\lceil dx\rceil}{d}, \frac{\lceil dy\rceil}{d}\Big) - A\Big(x, \frac{\lceil dy\rceil}{d}\Big)\Big| + \Big|A\Big(x, \frac{\lceil dy\rceil}{d}\Big) - A(x,y)\Big| \\ \le &\frac{2C_3}{d}. \end{aligned}$$

因此，对所有 $x,y\in[0,1]$，$|W_{\Sigma_d}(x,y) - A(x,y)| \le 2C_3 d^{-1}$。所声称的收敛结果由 $\|\cdot\|_\infty$ 的定义立即得出。类似地，我们可以将定义 2.3 推广到任意 m 阶张量，并研究定义在 $[0,1]^m$ 上且具有任意 $m\in\mathbb{N}$ 的函数的一致收敛。这样，我们可以将定义 1.2 中的 B 和 E 分别解释为对应于 x^t 的四阶矩张量和八阶矩张量的函数的 $C([0,1]^4)$-极限和 $C([0,1]^8)$-极限。

现在我们给出一个满足假设 1.2 的模型的具体示例。回顾一下，我们考虑的数据 $\{x_i^t\}_{t\ge1,\ i\in[d]}$ 是随机过程$\{W(s,\ x)\}_{s\in[0,\tau],\ x\in[0,1]}$ 的离散化，如定义 1.2 中所述。

正弦协方差。对于任意 $s\ge0$，令 $\{W(s,\ x)\}_{x\in[0,1]}$为一个中心化随机过程，其协方差函数 A 定义为

$$A(x,y) = a_0 + \sum_{k=1}^{\infty} b_k \cos\left(2\pi k\left(x-y\right)\right), \qquad \forall\ x,y\in[0,1]. \tag{18}$$

使得存在 $\varepsilon>0$ 使得

$$\sum_{k=1}^{\infty} k^{5+\varepsilon} b_k^2 < \infty. \tag{19}$$

定义有限维数据协方差矩阵为

$$\Sigma_d(i,j) = a_0 + \sum_{k=1}^{\infty} b_k \cos\left(2\pi k(i-j)/d\right), \qquad \forall\, i,j \in [d]. \tag{20}$$

令 A 如（18）中所定义，且 W_{Σ_d} 如定义2.3中所定义。当（19）成立时，很容易看出，当 $d \to \infty$ 时，$\|A - W_{\Sigma_d}\|_\infty \to 0$。条件（19保证（ A 在（18）中定义，并满足假设1.2。

若进一步假设 $W(s,\ \cdot)$ 是高斯过程，那么借助伊瑟利斯定理，四阶矩和八阶矩的表达式相对容易得到。也就是说，对于任意 $k \in \mathbb{N}$，有

$$\mathbb{E}\big[\prod_{i=1}^{2k} W(s,x_i)\big] = \sum_{p \in P_{2k}^2} \prod_{(i,j)\in p} \mathbb{E}\big[W(s,x_i)W(s,x_j)\big], \tag{21}$$

其中 P_{2k}^2 是 $\{1,\cdots,2k\}$ 的所有配对集合。基于此，我们得到如下结果 lt.

引理 2.4。 由 A 根据 (18) 和 (19) 定义的函数满足假设*1.2*。此外，若W 为高斯随机过程，则其四阶矩 B 和八阶矩 E 满足假设*1.2*。

我们在此强调，引理 2.4 中的高斯假设纯粹是为了简化证明。基于伊瑟利斯定理的推广，可以在椭圆分布情形（Zuoetal.，2021年）、高斯混合情形（Vignat，2012年）以及更一般的情形下，进行类似于证明引理 2.4 所需的计算。

接下来，我们将更明确地说明这些尺度极限的衰减性质；它们决定了低噪声设定下有限维、有限迭代均方误差和预测误差的极限行为。

命题 2.5. 采用定理*1.6*中的假设，并考虑低噪声设定 .

(a) 对所有 $\tau \in (0,\infty)$，均有 预测误差(τ) $\le$均方误差$(0)/\alpha\tau$。

(b) 若进一步假设，对于连续函数空间 $L^2([0,1])$ 的标准正交基 $\{\phi_i\}_{i=1}^{\infty}$， $A\colon [0,1]^2 \to \mathbb{R}$ 可分解为 $A(x,y) = \sum_{i=1}^{K} \lambda_i \phi_i(x)\phi_i(y)$，其中该级数在 $L^2([0,1]^2)$ 中收敛，且 $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_K$，满足 $\lambda := \lambda_K > 0$，则对所有 $\tau \in (0,\infty)$，均有均方误差(τ) $\le$均方误差(0) $\exp(-2\alpha\lambda\tau)$。

下面给出两个具体示例，以说明上述命题。

（1）设协方差矩阵定义为

$$\Sigma_d(i,j) = 1 + \cos\left(2\pi(i-j)/d\right) \quad \text{for all} \quad i,j \in [d].$$

则，对于所有 $x,y \in [0,1]$，函数 A 由 $A(x,y) = 1 + \cos(2\pi(x-y))$ 给出。注意，A 也可以表示为

$$A(x,y) = 1 + \frac{1}{2}\big(\sqrt{2}\cos(2\pi x)\big)\big(\sqrt{2}\cos(2\pi y)\big) + \frac{1}{2}\big(\sqrt{2}\sin(2\pi x)\big)\big(\sqrt{2}\sin(2\pi y)\big).$$

因此 A 满足命题 2.5 （b）部分中的条件，其中 $\lambda_1 = 1$， $\lambda_2 = \lambda_3 = 1/2$。因此，我们考察按（均方误差 式（17）定义的均方误差。根据命题2.5，均方误差的标度极限 均方误差 满足 均方误差(τ) $\le$均方误差$(0)\exp(-\alpha\tau)$ 对任意 $\tau \in (0,\infty)$。

(2) 现在，设协方差矩阵为

$$\Sigma_d(i,j) = \Big(\frac{|i-j|}{d} - \frac{1}{2}\Big)^2 - 2\Big(\frac{|i-j|}{d} - \frac{1}{2}\Big)^4 \quad \text{for all} \quad i,j \in [d].$$

于是函数 A 由 $A(x,y) = \big(|x-y| - \frac{1}{2}\big)^2 - 2\big(|x-y| = \frac{1}{2}\big)^4$ 给出，对所有 $x,y \in [0,1]$ 均成立。通过傅里叶级数展开，我们还可得到

$$A(x,y) = a_0 + \sum_{k=1}^{\infty} \frac{b_k}{2}\big(\sqrt{2}\cos(2\pi x)\big)\big(\sqrt{2}\cos(2\pi y)\big) + \sum_{k=1}^{\infty} \frac{b_k}{2}\big(\sqrt{2}\sin(2\pi x)\big)\big(\sqrt{2}\sin(2\pi y)\big),$$

其中 $a_0 = 7/120$ 且 $b_k = 6/\pi^2 k^4$ 对所有 $k \geq 1$ 均成立。还可以验证，(19) 在 $\varepsilon = 1$ 的条件下成立。由于特征值不存在统一的正下界，我们考察 (17) 中定义的 预测误差 （PE）。根据命题 2.5 的 (a) 部分，对于任意 $\tau \in (0,\infty)$，$\mathsf{PE}(s) \leq$ 均方误差$(0)/\alpha\tau$。

更一般地，为了计算标度极限和涨落，尤其是在中等噪声和高噪声设定下，通常都需要借助数值方法。有关细节，例如可参见 Lord 等人 (2014年)，对此的全面研究则留待未来工作。

3 证明概述

我们证明背后的总体思路是，将（3）中的插值随机梯度下降迭代/涨落视为取值于空间 $C([0,\tau]; C([0,1]))$ 的随机变量序列来研究。在适当的假设下，我们首先证明这些随机变量序列在 $C([0,\tau]; C([0,1]))$ 中是紧的；这构成了我们分析的主要部分。因此，我们可以将算法中的不同项视为随机变量序列，并且每个子列都有一个进一步的子列，它与 $C([0,\tau]; C([0,1]))$ 中的序贯弱极限联合依分布收敛$^{\otimes m}$(m 这里算法中的项数为该数值。

接下来，我们通过考察每一项的序贯弱极限，确定联合序贯弱极限。借助斯科罗霍德表示定理，联合序贯弱极限可以表示为某个常微分方程/随机微分方程的解，而该解等于插值随机梯度下降迭代/涨落的序贯弱极限（依分布）。最后，我们通过证明常微分方程/随机微分方程的逐路径解具有唯一性，证明所有序贯弱极限都相同，因此，插值随机梯度下降迭代/涨落序列弱收敛到唯一的极限。为方便读者，我们在下文陈述斯科罗霍德表示定理。

定理 3.1 （斯科罗霍德表示定理（Billingsley （2009）中的定理 6.7））。 设$\{\mathbb{P}_n\}_{n\geq 1}$ 和 $\mathbb{P}$ 是定义在 $\mathbb{R}^d$ 上的概率测度（该空间赋予其 *Borel* σ-代数），且 $\{\mathbb{P}_n\}_{n\geq 1}$弱收敛于 $\mathbb{P}$。那么存在一个概率空间（Ω $\mathcal{F}$， $\mathbb{P}$），其上定义了 $\mathbb{R}^d$取值于 $\{X_n\}_{n\geq 1}$ 的随机变量和 X ，其分布分别为 $\{\mathbb{P}_n\}_{n\geq 1}$ 和 $\mathbb{P}$ ，使得 X_n几乎处处收敛于 X 。

3.1 定理1.6

我们的第一个结果证明了 $\{\Theta^{d,T}(\cdot,\ \cdot)\}_{d\geq 1,\ T>0}$在空间 $C([0,\tau];\ C([0,1]))$ 中的紧性。该结果基于两个技术性结果（见命题 4.1 和命题 4.2），这两个结果构成了我们分析的主要部分。

定理 *3.2*（紧性）。设假设 *1.1* 和 *1.2* 成立，且满足定理 *1.6* 中的初始条件。此外，假设存在一致常数 $C_{s,1} > 0$ ，使得 $\max(\eta dT, \eta\sigma T^{\frac{1}{2}}) \le C_{s,1}$。则在 *(3)* 中定义的 $\{\Theta^{d,T}\}_{d\ge 1,T>0}$ 在 $C([0,\tau]; C([0,1]))$ 中是紧的。因此，$\{\Theta^{d,T}\}_{d\ge 1,T>0}$ 的任意子列都存在一个进一步的弱收敛子列，其极限位于 $C([0,\tau]; C([0,1]))$ 中。

定理 3.3（极限识别）。应用定理 *3.2*中的条件。则对于任意 $\tau > 0$以及 $\{\Theta^{d,T}\}_{d\ge 1,T>0}$的任意子列 $\{\Theta^{d_k,T_k}\}_{k\ge 1}$，还存在一个子列，使其弱收敛至函数 $\Theta \in C([0,\tau]; C([0,1]))$，当 $d_k, T_k \to \infty$ 时成立。此外，对于任意有界光滑函数 $f\colon [0,\tau] \to \mathbb{R}$，以及任意 $s \in [0,\tau]$, $x \in [0,1]$，我们有

$$\begin{aligned} &-\int_0^s \Theta(u,x) f'(u)\mathrm{d}u + f(s)\Theta(s,x) - f(0)\Theta(0,x) \\ &= -\eta dT\Big(\int_0^s \int_0^1 f(u)A(x,y)\Theta(u,y)\mathrm{d}y\mathrm{d}u + o(1)\Big) + \sigma^2\eta^2 T\Big(\int_0^s f(u)\mathrm{d}\xi_1(u,x) + o(1)\Big). \end{aligned}$$

其中 $\{\xi_1(s,x)\}_{s\in[0,\tau],x\in[0,1]}$ 是一个高斯场过程，其协方差由 (7)给出。

上述 定理在第4.2节中得到证明。作为一个直接推论，我们有：

(a) 当 $\eta dT \to \alpha$ 和 $\sigma^2/d^2T \to 0$ 在 $d, T \to \infty$时，我们得到

$$-\int_0^s \Theta(u,x) f'(u)\mathrm{d}u + f(s)\Theta(s,x) - f(0)\Theta(0,x) = -\alpha\int_0^s \int_0^1 f(u)A(x,y)\Theta(u,y)\mathrm{d}y\mathrm{d}u. \quad (22)$$

(b) 当 $\eta dT \to \alpha$ 和 $\sigma^2/d^2T \to \beta^2$ 在 $d, T \to \infty$ 时，我们得到

$$\begin{aligned} &-\int_0^s \Theta(u,x) f'(u)\mathrm{d}u + f(s)\Theta(s,x) - f(0)\Theta(0,x) \\ &= -\alpha\int_0^s \int_0^1 f(u)A(x,y)\Theta(u,y)\mathrm{d}y\mathrm{d}u + \alpha^2\beta^2\int_0^s f(u)\mathrm{d}\xi_1(u,x). \end{aligned}$$

(c) 当 $\eta\sigma T^{\frac{1}{2}} \to \alpha$ 和 $\sigma^2/d^2T \to \infty$ 在 $d, T \to \infty$ 时，我们得到

$$-\int_0^s \Theta(u,x) f'(u)\mathrm{d}u + f(s)\Theta(s,x) - f(0)\Theta(0,x) = \alpha^2\int_0^s f(u)\mathrm{d}\xi_1(u,x).$$

我们的主要结果 定理 1.6 由上述结果推出；证明见第4.2节 .

3.2 定理1.10

接下来我们证明涨落结果。根据第 1.1 节中插值过程的定义，我们立即得到 $U^{d,T} \in C([0,\tau]; C([0,1]))$。推导 $\{U^{d,T}\}_{d\ge 1,T>0}$ 的极限的思路与推导 $\{\Theta^{d,T}\}_{d\ge 1,T>0}$ 的极限的思路相似。具体而言，我们首先得到以下紧性结果。

定理 *3.4*（紧性）。设假设 *1.1* 和 *1.2* 成立，并进一步假设满足定理 *1.6* 和定理 *1.10* 中的初始条件。若 $\eta = \frac{\alpha}{dT}$, $\sigma = o(dT^{\frac{1}{2}})$ 且存在一致常数 $C_{s,2}$，使得对所有 $d \ge 1, T > 0$, $\max(\gamma T^{-\frac{1}{2}}, \gamma d^{-1}, \gamma\sigma d^{-1}T^{-\frac{1}{2}}) \le C_{s,2}$，则按式 *(11)* 定义的 $\{U^{d,T}(\cdot,\cdot)\}_{d\ge 1,T>0}$ 在 $C([0,\tau]; C([0,1]))$ 中是紧的。因此，$\{U^{d,T}(\cdot,\cdot)\}_{d\ge 1,T>0}$ 的任意子列都存在一个进一步的弱收敛子列，其极限属于 $C([0,\tau]; C([0,1]))$。

定理 3.5（极限识别）。 应用定理 3.4 中的条件。对于任意 $\tau > 0$ 以及 $\{U^{d,T}\}_{d\geq 1}$ 的任意子列 $\{U^{d_k,T_k}\}_{k\geq 1,T>0}$，都存在一个进一步的子列弱收敛至随机函数 $U \in C([0,\tau];C([0,1]))$，其中 d_k, $T_k \to \infty$。此外，对于任意有界光滑函数 $f\colon [0,\tau]\to\mathbb{R}$，若存在一些一致非负常数 ζ,β，使得 $\gamma\sigma d^{-1}T^{-\frac{1}{2}}\to\beta$ 且 $\gamma T^{-\frac{1}{2}}\to\zeta$, U 满足：对于任意 $s\in[0,\tau], x\in[0,1]$,

$$
\begin{aligned}
&-\int_0^s U(u,x)f'(u)\mathrm{d}u + f(s)U(s,x) - f(0)U(0,x)\\
=&-\alpha\int_0^s\int_0^1 f(u)A(x,y)U(u,y)\mathrm{d}y\mathrm{d}u + \alpha\beta\int_0^s f(u)\mathrm{d}\xi_2(u,x)\\
&+\alpha\zeta\int_0^s f(u)\mathrm{d}\xi_3(u,x) + O(\gamma d^{-1}+\gamma T^{-1}) + o(1).
\end{aligned}
$$

上述定理在第5.2节证明。由此可立即得到：

(a) 当 $T = o(d^2)$、 $\sigma = O(d)$ 且 $\gamma T^{-\frac{1}{2}}\to\zeta\in(0,\infty)$ 时，可以证明 $O(\gamma d^{-1}+\gamma T^{-1}) = o(1)$。因此，我们得到

$$
\begin{aligned}
&-\alpha^{-1}\Big(\int_0^s U(u,x)f'(u)\mathrm{d}u + f(s)U(s,x) - f(0)U(0,x)\Big)\\
=&-\int_0^s\int_0^1 f(u)A(x,y)U(u,y)\mathrm{d}y\mathrm{d}u + \beta\int_0^s f(u)\mathrm{d}\xi_2(u,x)\\
&+\zeta\int_0^s f(u)\mathrm{d}\xi_3(u,x).
\end{aligned}
$$

(b) 当 $\max(d,T^{\frac{1}{2}}) \ll \sigma \ll dT^{\frac{1}{2}}$ 且 $\gamma\sigma d^{-1}T^{-\frac{1}{2}}\to\beta\in(0,\infty)$ 时，可以证明 $\gamma T^{-\frac{1}{2}}\to 0$ 且 $O(\gamma d^{-1}+\gamma T^{-1}) = o(1)$。因此，我们得到

$$
\begin{aligned}
&-\alpha^{-1}\Big(\int_0^s U(u,x)f'(u)\mathrm{d}u + f(s)U(s,x) - f(0)U(0,x)\Big)\\
=&-\int_0^s\int_0^1 f(u)A(x,y)U(u,y)\mathrm{d}y\mathrm{d}u + \beta\int_0^s f(u)\mathrm{d}\xi_2(u,x).
\end{aligned}
$$

(c) 当 $d = O(T^{\frac{1}{2}})$、$\sigma = O(T^{\frac{1}{2}})$ 且 $\gamma d^{-1}\to 0$ 时，可以证明 $\gamma\sigma d^{-1}T^{-\frac{1}{2}}\to 0$、$\gamma T^{-\frac{1}{2}}\to 0$ 且 $O(\gamma d^{-1}+\gamma T^{-1}) = o(1)$。因此，我们得到

$$
\begin{aligned}
&-\int_0^s U(u,x)f'(u)\mathrm{d}u + f(s)U(s,x) - f(0)U(0,x)\\
=&-\alpha\int_0^s\int_0^1 f(u)A(x,y)U(u,y)\mathrm{d}y\mathrm{d}u.
\end{aligned}
\tag{23}
$$

我们的主要结果定理1.10和备注1.12 由上述结果推出，其证明见第5节。

3.3 解的存在性与唯一性

为完整证明定理1.6 和1.10中的主要结果，我们还需要给出相应随机微分方程解的存在唯一性结果。常微分方程情形可由 Picard-Lindel¨of 定理得出（例如，参见Arnold(1992年)）。此外，由于（9）、（10）、（15）和（16）可视为（14）的某些退化形式，我们重点关注（14）解的存在唯一性。下文的命题3.6 对于（9）、（10）、（15）和（16）也成立。

定理 3.6。 假设*1.2* 成立。

1. 若方程 *(14)* 的初始条件满足 $U(0,\cdot) \in C([0,1])$，则存在唯一解$\{U(s,x)\}_{s\in[0,\tau],x\in[0,1]} \in C([0,\tau]; C([0,1]))$，使其满足随机微分方程 *(14)*。*2.* 若方程 *(14)* 的初始条件满足 $U(0,\cdot) \in L^2([0,1])$，则存在唯一解$\{U(s,x)\}_{s\in[0,\tau],x\in[0,1]} \in C([0,\tau]; L^2([0,1]))$，使其满足随机微分方程 *(14)*。此外，该解满足以下稳定性性质：

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\in[0,\tau]}\|U(s,\cdot)\|^2_{L^2([0,1])}\Big] \le 4\Big(&\mathbb{E}\big[\|U(0,\cdot)\|^2_{L^2([0,1])}\big] + C_2\alpha^2\beta^2\tau \\ &+ (C_5+C_2^2)\alpha^2\zeta^2\int_0^\tau \|\Theta(s,\cdot)\|^2_{L^2([0,1])}\,\mathrm{d}s\Big)\exp\left(4C_2^2\alpha^2\tau^2\right).\end{aligned} \tag{24}$$

我们指出，上述结果不仅提供了理论支持，还为数值计算解提供了可证明的必要条件；例如，详见 Lord 等人(2014年) 的详细讨论。

4 缩放极限的证明

4.1 $\{\Theta^{d,T}\}_{d\ge1,\ T>0}$的紧性

我们首先给出以下两个主要命题。

命题 4.1。 假设条件 *1.1* 和假设条件 *1.2* 成立，并且存在一个一致常数 $C_{s,\ 1}>0$，使得对于极限序列上的所有 $d\ge1$、$T>0$，均有 $\max(\eta dT、\ \eta\sigma T^{\frac{1}{2}})\le C_{s,\ 1}$。则存在一个正数常数 C，其取决于 τ 和 $C_{s,\ 1}$，使得对于任意 d、$T>0$ 、$i\in[d]$ 以及任意 $0\le t_1<t_2\le\lfloor\tau T\rfloor$,

$$\mathbb{E}\left[\left|\Delta\theta_i^{t_2}-\Delta\theta_i^{t_1}\right|^4\right] \le C\left(\frac{t_2-t_1}{T}\right)^2. \tag{25}$$

命题的证明 *4.1*。 对于任意 $0\le t_1<t_2\le\lfloor\tau T\rfloor:=N$，有

$$\begin{aligned}&\Delta\theta_i^{t_2}-\Delta\theta_i^{t_1}\\ =&-\eta\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d}\mathbb{E}[x_i^tx_j^t]\Delta\theta_j^t-\eta\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d}\left(x_i^tx_j^t-\mathbb{E}[x_i^tx_j^t]\right)\Delta\theta_j^t+\eta\sum_{t=t_1}^{t_2-1}x_i^t\varepsilon^t,\end{aligned}$$

这意味着

$$\begin{aligned}\left|\Delta\theta_i^{t_2}-\Delta\theta_i^{t_1}\right|^4 \le 27\eta^4 \underbrace{\Big(\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d}\mathbb{E}[x_i^t x_j^t]\Delta\theta_j^t\Big)^4}_{S_{i,1}} \\ +27\eta^4\underbrace{\Big(\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d}\left(x_i^t x_j^t-\mathbb{E}[x_i^t x_j^t]\right)\Delta\theta_j^t\Big)^4}_{S_{i,2}}+27\eta^4\underbrace{\Big(\sum_{t=t_1}^{t_2-1}x_i^t\varepsilon^t\Big)^4}_{S_{i,3}}.\end{aligned} \tag{26}$$

对空间指标 i 求和并取期望，得到

$$\frac{1}{d}\sum_{i=1}^{d}\mathbb{E}\big[|\Delta\theta_i^{t_2}-\Delta\theta_i^{t_1}|^4\big] \le \frac{27\eta^4}{d}\sum_{i=1}^{d}\mathbb{E}\big[S_{i,1}\big]+\frac{27\eta^4}{d}\sum_{i=1}^{d}\mathbb{E}\big[S_{i,2}\big]+\frac{27\eta^4}{d}\sum_{i=1}^{d}\mathbb{E}\big[S_{i,3}\big]. \tag{27}$$

接下来，我们将逐项估计（27）右端项。定义

$$m_t^d := \frac{1}{d}\sum_{i=1}^{d}\mathbb{E}\left[|\Delta\theta_i^t|^4\right].$$

首先有

$$\begin{aligned}\mathbb{E}\big[S_{i,1}\big] &= \mathbb{E}\big[\sum_{r_1,r_2,r_3,r_4=t_1}^{t_2-1}\sum_{j_1,j_2,j_3,j_4=1}^{d}\prod_{k=1}^{4}A(\frac{i}{d},\frac{j_k}{d})\prod_{k=1}^{4}\Delta\theta_{j_k}^{r_k}\big] \\ &\le C_2^4\mathbb{E}\big[\sum_{r_1,r_2,r_3,r_4}\sum_{j_1,j_2,j_3,j_4}\prod_{k=1}^{4}|\Delta\theta_{j_k}^{r_k}|\big] \\ &\le C_2^4 d^4(t_2-t_1)^3\sum_{t=t_1}^{t_2-1}m_t^d,\end{aligned}$$

其中，第一个不等式由假设 1.2 得出，最后一个不等式则由 $abcd \le \frac{1}{4}(a^4+b^4+c^4+d^4)$ 得出，对任意 a，b，c，$d\in\mathbb{R}$ 均成立。我们还有

$$\mathbb{E}\big[S_{i,2}\big] = \mathbb{E}\big[\sum_{r_1,r_2,r_3,r_4=t_1}^{t_2-1}\sum_{j_1,j_2,j_3,j_4=1}^{d}\prod_{k=1}^{4}\left(x_i^{r_k}x_{j_k}^{r_k}-\mathbb{E}[x_i^{r_k}x_{j_k}^{r_k}]\right)\prod_{k=1}^{4}\Delta\theta_{j_k}^{r_k}\big].$$

由于 x^t 与 $x^{t'}$ 在 $t\neq t'$ 时相互独立，上述求和中的大多数项均为零。剩下的项分为两类。为继续推导，令 $\mathcal{F}_k^d$ 为由 $\big\{\{x_i^t\}_{i\in[d],0\le t\le k},\{\varepsilon^t\}_{0\le t\le k},\{\Delta\theta_i^0\}_{i\in[d]}\big\}$ 生成的 σ-代数。

类型 1：当 $r_1=r_2=r_3=r_4$ 时，非零项的形式为

$$\sum_{j_1,j_2,j_3,j_4=1}^{d}\mathbb{E}\big[\prod_{k=1}^{4}\left(x_i^t x_{j_k}^t-\mathbb{E}\left[x_i^t x_{j_k}^t\right]\right)\prod_{k=1}^{4}\Delta\theta_{j_k}^t\big],$$

并且有 $3(t_2-t_1)$ 个这样的项。这些项的和可上界为

$$3\sum_{t=t_1}^{t_2-1}\sum_{j_1,j_2,j_3,j_4=1}^{d}\mathbb{E}\big[\prod_{k=1}^{4}\left(x_i^t x_{j_k}^t-\mathbb{E}\left[x_i^t x_{j_k}^t\right]\right)\prod_{k=1}^{4}\Delta\theta_{j_k}^t\big]$$

$$= 3\sum_{t=t_1}^{t_2-1}\sum_{j_1,j_2,j_3,j_4=1}^{d}\mathbb{E}\Big[\mathbb{E}[\prod_{k=1}^{4}\left(x_i^t x_{j_k}^t - \mathbb{E}\left[x_i^t x_{j_k}^t\right]\right)\prod_{k=1}^{4}\Delta\theta_{j_k}^t]|\mathcal{F}_{t-1}^d\Big]$$
$$\leq 3C_8'\sum_{t=t_1}^{t_2-1}\sum_{j_1,j_2,j_3,j_4=1}^{d}\mathbb{E}[\prod_{k=1}^{4}|\Delta\theta_{j_k}^t|] \leq C_8' d^4\sum_{t=t_1}^{t_2-1} m_t^d,$$

其中，第一个不等式由假设1.2和 $C_8' = 16C_8 + 32C_2^2C_5 + 16C_2^4$得出。最后一个不等式由$abcd \leq \frac{1}{4}(a^4+b^4+c^4+d^4)$ 对任意 a， b， c， $d \in \mathbb{R}$。

T类型 2：当 r_1, r_2, r_3, r_4 两两相等时，非零项的形式为

$$\sum_{j_1,j_2,j_3,j_4=1}^{d}\mathbb{E}[\prod_{k=1}^{2}\left(x_i^t x_{j_k}^t - \mathbb{E}[x_i^t x_{j_k}^t]\right)\prod_{k=3}^{4}\left(x_i^{t'} x_{j_k}^{t'} - \mathbb{E}[x_i^{t'} x_{j_k}^{t'}]\right)\prod_{k=1}^{2}\Delta\theta_{j_k}^t\prod_{k=3}^{4}\Delta\theta_{j_k}^{t'}]$$

其中 $t \neq t'$，这样的项共有 $3(t_2-t_1)^2 - 3(t_2-t_1)$ 个。为简化记号，对所有 $i, j \in [d]$ 和 $t \leq \lfloor \tau T \rfloor$，记 $\Delta X_{ij}^t := x_i^t x_j^t - \mathbb{E}[x_i^t x_j^t]$。那么，这些项的和可以上界估计为

$$3\sum_{t\neq t'}\sum_{j_1,j_2,j_3,j_4=1}^{d}\mathbb{E}[\prod_{k=1}^{2}\Delta X_{ij_k}^t\prod_{k=3}^{4}\Delta X_{ij_k}^{t'}\prod_{k=1}^{2}\Delta\theta_{j_k}^t\prod_{k=3}^{4}\Delta\theta_{j_k}^{t'}]$$
$$= 3\sum_{t\neq t'}\sum_{j_1,j_2,j_3,j_4=1}^{d}\mathbb{E}\Big[\mathbb{E}[\prod_{k=1}^{2}\Delta X_{ij_k}^t\prod_{k=3}^{4}\Delta X_{ij_k}^{t'}\prod_{k=1}^{2}\Delta\theta_{j_k}^t\prod_{k=3}^{4}\Delta\theta_{j_k}^{t'}]|\mathcal{F}_{t\vee t'-1}^d\Big]$$
$$\leq 3C_5'^2\sum_{t\neq t'}\sum_{j_1,j_2,j_3,j_4=1}^{d}\mathbb{E}[\prod_{k=1}^{2}|\Delta\theta_{j_k}^t|\prod_{k=3}^{4}|\Delta\theta_{j_k}^{t'}|] \leq 3C_5'^2 d^4(t_2-t_1)\sum_{t=t_1}^{t_2-1} m_t^d,$$

其中，第一个不等式由假设 1.2和 $C_5' = C_5 + C_2^2$得到。最后一个不等式由以下事实得到：对任意 $a, b, c, d \in \mathbb{R}$，有 $4(abcd) \leq (a^4+b^4+c^4+d^4)$。结合这两个上界，我们得到

$$\mathbb{E}[S_{i,2}] \leq 3\big(C_8' + C_5'^2(t_2-t_1)\big)d^4\sum_{t=t_1}^{t_2-1} m_t^d \leq \big(C_8' + C_5'^2\big)(t_2-t_1)d^4\sum_{t=t_1}^{t_2-1} m_t^d.$$

最后，由于假设 1.2，我们估计 $\mathbb{E}[S_{i,3}]$如下：

$$\mathbb{E}[S_{i,3}] = \sum_{t=t_1}^{t_2-1} B(\frac{i}{d},\frac{i}{d},\frac{i}{d},\frac{i}{d})\mathbb{E}\big[(\varepsilon^t)^4\big] + \sum_{t\neq t'} A(\frac{i}{d},\frac{i}{d})A(\frac{i}{d},\frac{i}{d})\mathbb{E}\left[(\varepsilon^t)^2\right]\mathbb{E}\left[(\varepsilon^{t'})^2\right]$$
$$\leq C_5C_1\sigma^4(t_2-t_1) + C_2^2\sigma^4(t_2-t_1)^2 \leq \big(C_5C_1 + C_2^2\big)\sigma^4(t_2-t_1)^2.$$

将我们的估计 $\mathbb{E}[S_{i,,1}]$, $\mathbb{E}[S_{i,2}]$, $\mathbb{E}[S_{i,3}]$代入（27），即可得到

$$\frac{1}{d}\sum_{i=1}^{d}\mathbb{E}\big[|\Delta\theta_i^{t_2} - \Delta\theta_i^{t_1}|^4\big] \leq 27\eta^4 d^4\big[C_2^4(t_2-t_1)^3 + 3\big(C_8' + C_5'^2\big)(t_2-t_1)\big] \times \sum_{t=t_1}^{t_2-1} m_t^d + 27\big(C_5C_1 + C_2^2\big)\eta^4\sigma^4(t_2-t_1)^2. \tag{28}$$

在（ $t_1 = 0,\ t_2 = t \le N$ 中取28。对于任意 $t \le N$,

$$m_t^d \le 8m_0^d + 216\eta^4 d^4 \big[C_2^4 t^3 + 3\big(C_8' + C_5'^2\big)t\big] \sum_{k=0}^{t-1} m_k^d + 216\big(C_5C_1 + C_2^2\big)\eta^4\sigma^4 t^2$$
$$\le 8m_0^d + 216\big(C_2^4 + 3C_8' + 3C_5'^2\big)\eta^4 d^4 t^3 \sum_{k=0}^{t-1} m_k^d + 216\big(C_5C_1 + C_2^2\big)\eta^4\sigma^4 t^2.$$

根据离散 Gronwall 不等式，对于任意 $t \le N$，我们有:

$$\begin{aligned}
m_t^d &\le 216\big(C_5C_1 + C_2^2\big)\eta^4\sigma^4 t^2 \\
&\quad + 216^2\big(C_5C_1 + C_2^2\big)\eta^8 d^4\sigma^4 t^3 \sum_{k=0}^{t-1} k^2 \exp\big(216\big(C_2^4 + 3C_8' + 3C_5'^2\big)\eta^4 d^4 t^3 (t-k-1)\big) \\
&\le 216\big(C_5C_1 + C_2^2\big)\eta^4\sigma^4 t^2 \\
&\quad + 216^2\big(C_5C_1 + C_2^2\big)\eta^8 d^4\sigma^4 t^6 \exp\big(216\big(C_2^4 + 3C_8' + 3C_5'^2\big)\eta^4 d^4 t^4\big) \\
&\le 216\big(C_5C_1 + C_2^2\big)C_{s,1}^4\tau^2 \\
&\quad + 216^2\big(C_5C_1 + C_2^2\big)C_{s,1}^8\tau^6 \exp\big(216\big(C_2^4 + 3C_8' + 3C_5'^2\big)C_{s,1}^4\tau^4\big),
\end{aligned}$$

其中，最后一个不等式由 $\max(\eta dT, \eta\sigma T^{\frac{1}{2}}) \le C_{s,1}$ 以及 $N \le \tau T$ 得出。我们将 m_t^d 的上界简化为 $m_t^d \le C_\tau$，其中 C_τ 是与 d, T, σ 无关的某个正数。将 m_t^d 的上界代入 (26)，并结合对 $\mathbb{E}[S_{i,,1}]$, $\mathbb{E}[S_{i,2}]$, $\mathbb{E}[S_{i,3}]$ 的估计，我们得到

$$\begin{aligned}
&\mathbb{E}\left[|\Delta\theta_i^{t_2} - \Delta\theta_i^{t_1}|^2\right] \\
&\le 27\eta^4 d^4\big(C_2^4(t_2-t_1)^3 + 3(C_8' + C_5'^2)(t_2-t_1)\big)\sum_{t=t_1}^{t_2-1} m_t^d + 27(C_5C_1 + C_2^2)\eta^4\sigma^4(t_2-t_1)^2 \\
&\le 27C_\tau\big(C_2^4 + 3C_8' + 3C_5'^2\big)\eta^4 d^4(t_2-t_1)^4 + 27(C_5C_1 + C_2^2)\eta^4\sigma^4(t_2-t_1)^2 \\
&\le 27C_\tau\big(C_2^4 + 3C_8' + 3C_5'^2\big)C_{s,1}^4\left(\frac{t_2-t_1}{T}\right)^4 + 27(C_5C_1 + C_2^2)C_{s,1}^4\left(\frac{t_2-t_1}{T}\right)^2,
\end{aligned}$$

其中，最后一个不等式由 $\max(\eta dT, \eta\sigma T^{\frac{1}{2}}) \le C_{s,1}$ 得出。最后，由于 $\left(\frac{t_2-t_1}{T}\right)^2 \le \tau^2$，对任意 $0 \le t_1 < t_2 \le N$，(25) 得证。 ■

命题 4.2。 假设假设 *1.1* 和假设 *1.2* 成立，且存在一致常数 $C_{s,1} > 0$，使得沿极限序列对所有 $d \ge 1$, $T > 0$ 均有 $\max(\eta dT, \eta\sigma T^{\frac{1}{2}}) \le C_{s,1}$。那么，存在一个依赖于 τ 和 $C_{s,1}$ 的正数 C，使得对任意 d, $T > 0$, $i, j \in [d]$ 和 $0 \le t_1 < t_2 \le \lfloor \tau T \rfloor$,

$$\mathbb{E}\left[\left|(\Delta\theta_i^{t_2} - \Delta\theta_j^{t_2}) - (\Delta\theta_i^{t_1} - \Delta\theta_j^{t_1})\right|^4\right] \le C\left(\frac{i-j}{d}\right)^4\left(\frac{t_2-t_1}{T}\right)^2. \tag{29}$$

此外，若定理 *1.6* 中的初始条件成立，则对任意 i, $j \in [d]$ 以及任意 $0 \le t \le \lfloor \tau T \rfloor$,

$$\mathbb{E}\left[\left|\Delta\theta_i^t - \Delta\theta_j^t\right|^4\right] \le \left(8L^4 + 8\tau^2 C\right)\left(\frac{i-j}{d}\right)^4. \tag{30}$$

其中 L 是定理 *1.6* 中初始条件的一致常数。

命题 4.2 的证明。 根据（2），对于任意 $i,\ j \in [d]$ 以及任意 $0 \le t \le \lfloor \tau T \rfloor - 1 := N-1$，我们有

$$\begin{aligned}\Delta\theta_i^{t+1} - \Delta\theta_j^{t+1} = (\Delta\theta_i^t - \Delta\theta_j^t) - \eta \sum_{l=1}^{d} \left(\mathbb{E}\left[x_i^t x_l^t\right] - \mathbb{E}\left[x_j^t x_l^t\right]\right) \Delta\theta_l^t \\ - \eta \sum_{l=1}^{d} \left(x_i^t x_l^t - \mathbb{E}\left[x_i^t x_l^t\right] - x_j^t x_l^t + \mathbb{E}\left[x_j^t x_l^t\right]\right) \Delta\theta_l^t + \eta(x_i^t - x_j^t)\varepsilon^t.\end{aligned}$$

对时间指标求和，并对等式两边绝对值取期望后，对于任意 $0 \le t_1 < t_2 \le N$，我们得到

$$\begin{aligned}&\mathbb{E}\left[\left|(\Delta\theta_i^{t_2} - \Delta\theta_j^{t_2}) - (\Delta\theta_i^{t_1} - \Delta\theta_j^{t_1})\right|^4\right] \\ &\le 27\eta^4 \mathbb{E}\Big[\underbrace{\Big(\sum_{t=t_1}^{t_2-1} \sum_{l=1}^{d} \left(\mathbb{E}\left[x_i^t x_l^t\right] - \mathbb{E}\left[x_j^t x_l^t\right]\right) \Delta\theta_l^t\Big)^4}_{\mathbb{E}[S_{i,j,1}]}\Big] \\ &\quad + 27\eta^4 \mathbb{E}\Big[\underbrace{\Big(\sum_{t=t_1}^{t_2-1} \sum_{l=1}^{d} \left(x_i^t x_l^t - \mathbb{E}\left[x_i^t x_l^t\right] - x_j^t x_l^t + \mathbb{E}\left[x_j^t x_l^t\right]\right) \Delta\theta_l^t\Big)^4}_{\mathbb{E}[S_{i,j,2}]}\Big] \\ &\quad + 27\eta^4 \mathbb{E}\Big[\underbrace{\Big(\sum_{t=t_1}^{t_2-1} (x_i^t - x_j^t)\varepsilon^t\Big)^4}_{\mathbb{E}[S_{i,j,3}]}\Big].\end{aligned} \tag{31}$$

接下来，我们将逐项估计31式右端。首先有

$$\begin{aligned}\mathbb{E}\big[S_{i,j,1}\big] &= \mathbb{E}\Big[\sum_{r_1,r_2,r_3,r_4=t_1}^{t_2-1} \sum_{l_1,l_2,l_3,l_4=1}^{d} \prod_{k=1}^{4} \left(A(\frac{i}{d}, \frac{j_k}{d}) - A(\frac{j}{d}, \frac{j_k}{d})\right) \prod_{k=1}^{4} \Delta\theta_l^t\Big] \\ &\le C_3^4 \big(\frac{i-j}{d}\big)^4 (t_2 - t_1)^3 d^4 \sum_{t=t_1}^{t_2-1} m_t^d,\end{aligned}$$

其中，不等式由假设1.2 以及 $\mathbf{4}(abcd) \le a^4 + b^4 + c^4 + d^4$ 对所有 $a,\ b,\ c,\ d \in \mathbb{R}$成立得出。类似于 $\mathbb{E}[S_{i,2}]$ 命题4.1证明中的做法，在估计 $\mathbb{E}[S_{i,j,2}]$时，存在两类非零项。

类型 1：第一类非零项的形式为

$$\sum_{l_1,l_2,l_3,l_4=1}^{d} \mathbb{E}\Big[\prod_{k=1}^{4} \left(x_i^t x_{l_k}^t - \mathbb{E}[x_i^t x_{l_k}^t] - x_j^t x_{l_k}^t + \mathbb{E}[x_j^t x_{l_k}^t]\right) \Big| \prod_{k=1}^{4} \Delta\theta_{l_k}^t \Big|\Big],$$

且此类项共有 $\mathbf{3}(t_2 - t_1)$ 项。此类项之和可由以下内容界定：

$$3 \sum_{t=t_1}^{t_2-1} \sum_{l_1,l_2,l_3,l_4=1}^{d} \mathbb{E}\Big[\prod_{k=1}^{4} \left(x_i^t x_{l_k}^t - \mathbb{E}[x_i^t x_{l_k}^t] - x_j^t x_{l_k}^t + \mathbb{E}[x_j^t x_{l_k}^t]\right) \Big| \prod_{k=1}^{4} \Delta\theta_{l_k}^t \Big|\Big]$$

$$= 3\sum_{t=t_1}^{t_2-1}\sum_{l_1,l_2,l_3,l_4=1}^{d}\mathbb{E}\Big[\mathbb{E}\big[\prod_{k=1}^{4}\big(x_i^t x_{l_k}^t - \mathbb{E}[x_i^t x_{l_k}^t] - x_j^t x_{l_k}^t + \mathbb{E}[x_j^t x_{l_k}^t]\big) \,\big|\, \prod_{k=1}^{4}\Delta\theta_{l_k}^t\big|\big]\big|\mathcal{F}_{t-1}^d\Big]$$

$$\le 3C_9'^4\big(\frac{i-j}{d}\big)^4\sum_{t=t_1}^{t_2-1}\sum_{l_1,l_2,l_3,l_4=1}^{d}\mathbb{E}\big[\big|\prod_{k=1}^{4}\Delta\theta_{l_k}^t\big|\big] \le C_9'^4\big(\frac{i-j}{d}\big)^4 d^4\sum_{t=t_1}^{t_2-1} m_t^d,$$

其中，第一个不等式由假设 1.2 以及 $C_9'^4 = C_9^4 + 2C_3^2C_6^2 + C_3^4$ 得出。最后一个不等式由 $4(abcd) \le a^4+b^4+c^4+d^4$ 对所有 a, b, c, $d \in \mathbb{R}$ 得出。

类型 2：为简化记号，对于所有 $i,j \in [d]$ 和 $t \le \lfloor \tau T \rfloor$，我们记 $\Delta X_{ij}^t := x_i^t x_j^t - \mathbb{E}\big[x_i^t x_j^t\big]$。于是，第二类非零项的形式为

$$\sum_{l_1,l_2,l_3,l_4=1}^{d}\mathbb{E}\Bigg[\underbrace{\prod_{k=1}^{2}\big(\Delta X_{il_k}^t - \Delta X_{jl_k}^t\big)\prod_{k=3}^{4}\Big(\Delta X_{il_k}^{t'} - \Delta X_{jl_k}^{t'}\Big)}_{P_{i,j,l_1,l_2,l_3,l_4}^{t,t'}} \times \big|\prod_{k=1}^{2}\Delta\theta_{l_k}^t\prod_{k=3}^{4}\Delta\theta_{l_k}^{t'}\big|\Bigg],$$

其中 $t \ne t'$，且共有 $3(t_2-t_1)^2 - 3(t_2-t_1)$ 个这样的项。这些项的和可以上界为

$$\sum_{t\ne t'}\sum_{l_1,l_2,l_3,l_4=1}^{d}\mathbb{E}\Big[\mathbb{E}\big[P_{i,j,l_1,l_2,l_3,l_4}^{t,t'}\prod_{k=1}^{2}|\Delta\theta_{l_k}^t|\prod_{k=3}^{4}|\Delta\theta_{l_k}^{t'}|\big]\big|\mathcal{F}_{t\vee t'-1}^d\Big]$$

$$\le C_6'^4\big(\frac{i-j}{d}\big)^4\sum_{t\ne t'}\sum_{l_1,l_2,l_3,l_4=1}^{d}\mathbb{E}\big[\prod_{k=1}^{2}|\Delta\theta_{l_k}^t|\prod_{k=3}^{4}|\Delta\theta_{l_k}^{t'}|\big] \le C_6'^4\big(\frac{i-j}{d}\big)^4 d^4(t_2-t_1)\sum_{t=t_1}^{t_2-1} m_t^d,$$

其中，第一个不等式由假设 1.2 和 $C_6'^4 = C_3^2 + C_6^2$得出。最后一个不等式由 $abcd \le \frac{a^4+b^4+c^4+d^4}{4}$ 对所有 $a,b,c,d \in \mathbb{R}$得出。~~因此，~~ $\mathbb{E}[S_{i,j,2}]$ 可以上界为

$$\mathbb{E}[S_{i,j,2}] \le \big(C_9'^4 + C_6'^4(t_2-t_1)\big)\big(\frac{i-j}{d}\big)^4 d^4\sum_{t=t_1}^{t_2-1} m_t^d$$

$$\le \big(C_9'^4 + C_6'^4\big)(t_2-t_1)\big(\frac{i-j}{d}\big)^4 d^4\sum_{t=t_1}^{t_2-1} m_t^d.$$

根据假设 1.2，我们可以将 $\mathbb{E}[S_{i,j,3}]$界定为

$$\mathbb{E}[S_{i,j,3}]$$

$$= \sum_{t=t_1}^{t_2-1}\mathbb{E}\big[(x_i^t - x_j^t)^4\big]\mathbb{E}\big[\big(\varepsilon^t\big)^4\big] + \sum_{t\ne t'}\mathbb{E}\big[(x_i^t - x_j^t)^2\big]\mathbb{E}\big[(x_i^{t'} - x_j^{t'})^2\big]\mathbb{E}\big[\big(\varepsilon^t\big)^2\big]\mathbb{E}\big[\big(\varepsilon^{t'}\big)^2\big]$$

$$\le C_1C_7^4(t_2-t_1)\big(\frac{i-j}{d}\big)^4\sigma^4 + C_4^4(t_2-t_1)^2\big(\frac{i-j}{d}\big)^4\sigma^4$$

$$\le \big(C_1C_7^4 + C_4^4\big)(t_2-t_1)^2\big(\frac{i-j}{d}\big)^4\sigma^4.$$

我们已在命题 $m_t^d \le C_\tau$ 的证明中表明，对任意 $0 \le t \le N$，4.1均有。结合（31）以及对 $\mathbb{E}[S_{i,j,1}]$、$\mathbb{E}[S_{i,j,2}]$、$\mathbb{E}[S_{i,j,3}]$的估计，我们有

$$\mathbb{E}\big[\big|(\Delta\theta_i^{t_2} - \Delta\theta_j^{t_2}) - (\Delta\theta_i^{t_1} - \Delta\theta_j^{t_1})\big|^4\big]$$

$$
\begin{aligned}
&\leq 27(C_9'^4 + C_6'^4 + C_3^4)(t_2 - t_1)^3 \big(\frac{i-j}{d}\big)^4 \eta^4 d^4 \sum_{t=t_1}^{t_2-1} m_t^d \\
&\quad + 27\big(C_1 C_7^4 + C_4^4\big)(t_2 - t_1)^2 \big(\frac{i-j}{d}\big)^4 \eta^4 \sigma^4 \\
&\leq 27 C_\tau (C_9'^4 + C_6'^4 + C_3^4) C_{s,1}^4 \big(\frac{t_2 - t_1}{T}\big)^4 \big(\frac{i-j}{d}\big)^4 + 27 C_{s,1}^4 \big(\frac{t_2 - t_1}{T}\big)^2 \big(\frac{i-j}{d}\big)^4,
\end{aligned}
$$

其中，最后一个不等式由 $\max(\eta d T, \eta\sigma T^{\frac{1}{2}}) \leq C_{s,1}$推出。因此，（29）得证。最后，（30）由（29）以及定理1.6中的初始条件推出。■

现在，我们可以基于以上两个命题证明定理 3.2 。

定理 *3.2*的证明。 根据柯尔莫哥洛夫紧性判据可以证明紧性（Karatzas 和 Shreve， 2012年，第 4 章）。最后一个结论直接由紧性性质得出。为了应用柯尔莫哥洛夫紧性判据，我们需要验证以下两个条件:

(a) $\{\Theta^{d,T}(0,\ 0)\}_{d\geq 1,\ T>0}$ 在概率空间中是紧的。

(b) 存在一个正数常量 $C_{\mathbf{tight}}$，使得对于任意 $s_1,\ s_2 \in [0,\tau]$ 和 $x_1,\ x_2 \in [0,1]$，我们有

$$
\sup_{d,T} \mathbb{E}\big[\big|\overline{\Theta}^{d,T}(s_1,x_1) - \overline{\Theta}^{d,T}(s_2,x_2)\big|^4\big] \leq C_{\mathbf{tight}}\big(|s_1 - s_2|^2 + |x_1 - x_2|^4\big).
$$

为了验证 (*a*)，容易看出，对于任意 $d \geq 1,\ T > 0$ 和 $N > 0$,

$$
\mathbb{P}\big(\big|\overline{\Theta}^{d,T}(0,0)\big| > N\big) \leq N^{-2}\mathbb{E}\big[\big|\Delta\theta_0^0\big|^2\big] \leq N^{-2}R^2 \to 0 \qquad \text{as } N \to \infty.
$$

为了验证 (*b*)，首先注意到 $\Theta^{d,\ T}(\cdot,\cdot)$ 关于两个变量都是分段线性的，这是由 (3) 得出的。不失一般性，我们假设 $0 \leq s_1 < s_2 \leq \tau,\ \ 0 \leq x_1 < x_2 \leq 1$和 $\lfloor T s_1 \rfloor = t_1 < \lfloor T s_2 \rfloor = t_2$, $\lfloor d x_1 \rfloor = i < \lfloor d x_2 \rfloor = j$。我们有

$$
\begin{aligned}
\mathbb{E}\big[\big|\overline{\Theta}^{d,T}(s_1,x_1) - \overline{\Theta}^{d,T}(s_2,x_2)\big|^4\big] \leq\ & 8\mathbb{E}\big[\big|\overline{\Theta}^{d,T}(s_1,x_1) - \overline{\Theta}^{d,T}(s_2,x_1)\big|^4\big] \\
& + 8\mathbb{E}\big[\big|\overline{\Theta}^{d,T}(s_2,x_1) - \overline{\Theta}^{d,T}(s_2,x_2)\big|^4\big].
\end{aligned}
$$

根据命题4.1，我们有

$$
\begin{aligned}
&\mathbb{E}\big[\big|\overline{\Theta}^{d,T}(s_1,x_1) - \overline{\Theta}^{d,T}(s_2,x_1)\big|^4\big] \\
&\leq 27C\left(\big(\frac{t_1+1}{T} - s_1\big)^2 + \big(\frac{t_2 - t_1 - 1}{T}\big)^2 + \big(s_2 - \frac{t_2}{T}\big)^2\right) \\
&\leq 27C\,(s_2 - s_1)^2\,.
\end{aligned}
$$

根据命题4.2，我们有

$$
\begin{aligned}
&\mathbb{E}\big[\big|\overline{\Theta}^{d,T}(s_2,x_1) - \overline{\Theta}^{d,T}(s_2,x_2)\big|^4\big] \\
&\leq 27(8M^4 + 8\tau^2 C)\left(\big(\frac{i+1}{d} - x_1\big)^4 + \big(\frac{j-i-1}{d}\big)^4 + \big(x_2 - \frac{j}{d}\big)^4\right) \\
&\leq 216(M^4 + \tau^2 C)\,(x_2 - x_1)^4\,.
\end{aligned}
$$

因此，(*b*) 在 $C_{\mathbf{tight}} = 216C + 1728(M^4 + \tau^2 C)$ 下成立。 ■

4.2 极限识别

定理的证明 *3.3*。 根据定理 3.2，任意子列

$$\{\overline{\Theta}^{d_k,T_k}(\cdot,\cdot)\}_{k\geq 1} \quad \text{of} \quad \{\overline{\Theta}^{d,T}(\cdot,\cdot)\}_{d\geq 1,T>0}$$

存在一个进一步的弱收敛子列，其极限为 $\Theta \in C([0,\tau];\, C([0,1]))$，当 $d_k, T_k \to \infty$ 时。为简化记号，在证明中我们将 $\{\Theta^{d_k,T_k}\}_{k\geq 1}$ 的收敛子列记为 $\{\Theta^{d,T}\}_{d\geq 1,T>0}$。

为识别极限，首先我们用 $\Theta^{d,T}$ 重写（2）。对于任意 $0 \leq t \leq \lfloor \tau T \rfloor - 1$ 和任意 $i \in [d]$:

$$\begin{aligned} &\overline{\Theta}^{d,T}(\frac{t+1}{T}, \frac{i}{d}) - \overline{\Theta}^{d,T}(\frac{t}{T}, \frac{i}{d}) \\ =& -\eta \sum_{j=1}^{d} W(\frac{t}{T}, \frac{i}{d}) W(\frac{t}{T}, \frac{j}{d}) \overline{\Theta}^{d,T}(\frac{t}{T}, \frac{j}{d}) + \eta W(\frac{t}{T}, \frac{i}{d}) \varepsilon^t. \end{aligned} \tag{32}$$

因此，对于任意有界光滑函数 $f\colon\ [0,\tau] \to \mathbb{R}$，以及任意 $s \in [0,\tau]$、$i \in [d-1]$ 和 $x \in [\frac{i}{d}, \frac{i+1}{d}]$，我们有

$$\begin{aligned} &\sum_{t=0}^{\lfloor sT \rfloor - 1} f(\frac{t}{T}) \big(\overline{\Theta}^{d,T}(\frac{t+1}{T}, x) - \overline{\Theta}^{d,T}(\frac{t}{T}, x) \big) \\ =& -\eta(i+1-dx) \sum_{t=0}^{\lfloor sT \rfloor - 1} \sum_{j=1}^{d} f(\frac{t}{T}) W(\frac{t}{T}, \frac{i}{d}) W(\frac{t}{T}, \frac{j}{d}) \overline{\Theta}^{d,T}(\frac{t}{T}, \frac{j}{d}) \\ &+ \eta(i+1-dx) \sum_{t=0}^{\lfloor sT \rfloor - 1} f(\frac{t}{T}) W(\frac{t}{T}, \frac{i}{d}) \varepsilon^t \\ &- \eta(dx-i) \sum_{t=0}^{\lfloor sT \rfloor - 1} \sum_{j=1}^{d} f(\frac{t}{T}) W(\frac{t}{T}, \frac{i+1}{d}) W(\frac{t}{T}, \frac{j}{d}) \overline{\Theta}^{d,T}(\frac{t}{T}, \frac{j}{d}) \\ &+ \eta(dx-i) \sum_{t=0}^{\lfloor sT \rfloor - 1} f(\frac{t}{T}) W(\frac{t}{T}, \frac{i+1}{d}) \varepsilon^t. \end{aligned} \tag{33}$$

我们可以将（33）的左侧改写为

$$\begin{aligned} &\text{LHS}(33) \\ =& \sum_{t=1}^{\lfloor sT \rfloor - 1} \overline{\Theta}^{d,T}(\frac{t}{T}, x) \big(f(\frac{t-1}{T}) - f(\frac{t}{T}) \big) + f(\frac{\lfloor sT \rfloor - 1}{T}) \overline{\Theta}^{d,T}(\frac{\lfloor sT \rfloor}{T}, x) - f(0) \overline{\Theta}^{d,T}(0, x). \end{aligned}$$

当 $d, T \to \infty$ 时，由于 f 有界且光滑，对于任意 $s \in (0,\tau)$，我们有

$$f(s \pm T^{-1}) = f(s) + O(T^{-1}), \qquad T\left(f(s) - f(s - T^{-1})\right) = f'(s) + O(T^{-1}).$$

由于 $\{\overline{\Theta}^{d,T}\}_{d\geq 1,T>0}$ 弱收敛到 Θ，且 f, f' 连续有界，因此我们有

$$\frac{1}{T} \sum_{t=1}^{\lfloor sT \rfloor - 1} \overline{\Theta}^{d,T}(\frac{t}{T}, x) \left(f'(\frac{t}{T}) + O(T^{-1}) \right) = \int_0^{\frac{\lfloor sT \rfloor}{T}} \Theta(u,x) f'(u) \mathrm{d}u + o(1).$$

因此

$$\text{LHS}(33) = -\int_0^{\frac{\lfloor sT\rfloor}{T}} \Theta(u,x) f'(u)\mathrm{d}u + f(s)\Theta(\frac{\lfloor sT\rfloor}{T},x) - f(0)\Theta(0,x) + o(1). \tag{34}$$

接下来，我们考察 (33) 的右端。对于任意 $i \in [d]$，$\text{RHS}(33) = -(i+1-dx)I_i^s - (dx-i)I_{i+1}^s$，

$$\begin{aligned} & I_i^s \\ =& \underbrace{-\frac{\eta dT}{dT}\sum_{t=0}^{\lfloor sT\rfloor-1}\sum_{j=1}^{d} f(\frac{t}{T})A(\frac{i}{d},\frac{j}{d})\overline{\Theta}^{d,T}(\frac{t}{T},\frac{j}{d})}_{N_{i,1}^s} + \underbrace{\frac{\eta dT}{dT}\sum_{t=0}^{\lfloor sT\rfloor-1} f(\frac{t}{T})W(\frac{t}{T},\frac{i}{d})\varepsilon^t}_{N_{i,2}^s} \\ & \underbrace{-\frac{\eta dT}{dT}\sum_{t=0}^{\lfloor sT\rfloor-1}\sum_{j=1}^{d} f(\frac{t}{T})\big(W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d}) - \mathbb{E}\big[W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})\big]\big)\overline{\Theta}^{d,T}(\frac{t}{T},\frac{j}{d})}_{N_{i,3}^s}. \end{aligned}$$

基于第1节中引入的相同的时空插值，对于每个 $l = 1, 2, 3$，$\{N_{i,l}^s\}$ 都可以插值为 $C([0,\tau]; C([0,1])$ 中的一个函数，记为 $N^{d,T}_{i,l}$。根据引理 4.3，$(\Theta^{d,T}, N^{d,T}_{i,1}, N^{d,T}_{i,2}, N^{d,T}_{i,3})$ 在 $C([0,\tau]; C([0,1])^{\otimes 4}$ 中是紧的。因此，序贯弱极限存在于 $C([0,\tau]; C([0,1])^{\otimes 4}$ 中。$(\Theta^{d,T}, N^{d,T}_{i,1}, N^{d,T}_{i,2}, N^{d,T}_{i,3})$ 在 $C([0,\tau]; C([0,1])^{\otimes 4}$ 中的序贯弱极限，是基于各项的弱极限推导得到的。注意，对于 $\Theta^{d,T}$，其弱极限记为 Θ，满足一个基于其他项极限的方程，这些极限的确定如下：

- 对于 $N^{d,T}_{i,1}$，基于连续映射定理确定弱极限（Billingsley，2009，定理 2.7）；
- 对于 $N^{d,T}_{i,2}$，要确定无限维过程的弱极限，只需确定每个有限维弱极限的弱极限。我们通过逐点中心极限定理确定有限维极限。
- 对于 $N^{d,T}_{i,3}$我们利用这样一个事实：依概率收敛到零意味着依分布收敛到零。

我们从项 $N^{d,T}_{i,1}$ 开始。由于 f，A连续有界，且 $\{\Theta^{d,T}\}_{d\geq 1,\, T>0}$弱收敛于 Θ（见定理 3.2），因此有

$$\begin{aligned} N_{i,1}^s &= -\eta dT\big(\frac{1}{dT}\sum_{t=0}^{\lfloor sT\rfloor-1}\sum_{j=1}^{d} f(\frac{t}{T})A(\frac{i}{d},\frac{j}{d})\Theta(\frac{t}{T},\frac{j}{d}) + o(1)\big) \\ &= -\eta dT\big(\int_0^{\frac{\lfloor sT\rfloor}{T}}\int_0^1 f(u)A(\frac{i}{d},y)\Theta(u,y)\mathrm{d}y\mathrm{d}u + o(1)\big). \end{aligned}$$

对于 $N^s_{i,2} = \frac{\eta dT}{dT}\sum_{t=0}^{\lfloor sT\rfloor-1} f(\frac{t}{T})W(\frac{t}{T},\frac{i}{d})\varepsilon^t$，注意到 $\mathbb{E}[N^s_{i,2}] = 0$。由于 $\{W(\frac{t}{T},\frac{i}{d})\varepsilon^t\}_{t=1}^T$是一个独立同分布随机变量序列，因此可以利用标准中心极限定理研究 $N^s_{i,2}$ 的极限。特别地，我们有

$$\mathbb{E}[N_{i,2}^{s_1}N_{j,2}^{s_2}] = \sigma^2\eta^2 TA(\frac{i}{d},\frac{j}{d})\big(\int_0^{s_1\wedge s_2} f(u)^2\mathrm{d}u + o(1)\big). \tag{35}$$

因此，存在一个中心化高斯场过程 $\{\xi_1(s,\ x)\}_{s\in[0,\tau],\ x\in[0,1]}$，使得对于任意$s_1,\ s_2\in[0,\tau]$和 $x,\ y\in[0,1]$,

$$\mathbb{E}[\xi_1(s_1,x)\xi_1(s_2,y)]=\sigma_1(s_1,s_2,x,y)\quad\text{with}\quad\sigma_1(s_1,s_2,x,y)=(s_1\wedge s_2)A(x,y),$$

并且，对所有 $i\in[d]$, $N_{i,2}^s=\sigma^2\eta^2T\big(\int_0^s f(u)\mathrm{d}\xi_1(u,\frac{i}{d})+o(1)\big)$依分布收敛。通过方程（35）检验随机过程 $\{(N_{1,3}^s,N_{2,3}^s,\cdots,N_{d,3}^s)\}_{s\geq0}$的协方差，我们得到 $\{(N_{i,2}^s)_{i\in[d]}\}_{s\geq0}=\{(\sigma^2\eta^2T\int_0^s f(u)\mathrm{d}\xi_1(u,\frac{i}{d})+o(1))_{i\in[d]}\}_{s\geq0}$。对于$N_{i,3}^s$，我们有

$$N_{i,3}^s=-(\eta dT)\sum_{j=1}^{d}\sum_{t=0}^{\lfloor sT\rfloor-1}Z_{t,j},$$

其中$Z_{t,\ j}:=\frac{1}{dT}f(\frac{t}{T})\big(W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})-\mathbb{E}\big[W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})\big]\big)\Theta^d(\frac{t}{T},\frac{j}{d})$。注意，

$$\begin{aligned}&\mathbb{E}\left[Z_{t,j}\right]\\=&\frac{1}{T}f(\frac{t}{T})\mathbb{E}\bigg[\mathbb{E}\big[\big(W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})-\mathbb{E}\big[W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})\big]\big)|\mathcal{F}_{t-1}^d\big]\overline{\Theta}^{d,T}(\frac{t}{T},\frac{j}{d})\bigg]=0.\end{aligned}$$

此外，我们可以验证，对于任意 $t_1\neq t_2$ 以及任意 $j,\ l\in[d]$,

$$\begin{aligned}&\mathbb{E}\left[Z_{t_1,j}Z_{t_2,l}\right]\\=&\frac{1}{d^2T^2}f(\frac{t_1}{T})f(\frac{t_2}{T})\big(B(\frac{i}{d},\frac{j}{d},\frac{i}{d},\frac{l}{d})-A(\frac{i}{d},\frac{j}{d})A(\frac{i}{d},\frac{l}{d})\big)\mathbb{E}\big[\overline{\Theta}^{d,T}(\frac{t_1}{T},\frac{j}{d})\overline{\Theta}^{d,T}(\frac{t_2}{T},\frac{l}{d})\big].\end{aligned}$$

因此，我们有 $\mathbb{E}\left[N_{i,3}^s\right]=0$ 以及

$$\begin{aligned}&\mathrm{Var}(N_{i,3})\\=&\eta^2\sum_{j,l=1}^{d}\sum_{t=0}^{\lfloor sT\rfloor-1}f(\frac{t}{T})^2\big(B(\frac{i}{d},\frac{j}{d},\frac{i}{d},\frac{l}{d})-A(\frac{i}{d},\frac{j}{d})A(\frac{i}{d},\frac{l}{d})\big)\mathbb{E}\big[\overline{\Theta}^{d,T}(\frac{t}{T},\frac{j}{d})\overline{\Theta}^{d,T}(\frac{t}{T},\frac{l}{d})\big]\\\leq&\left(C_2^2+C_5\right)\|f\|_\infty^2\,s\eta^2dT\sum_{j=1}^{d}\mathbb{E}\big[|\overline{\Theta}^{d,T}(\frac{t}{T},\frac{j}{d})|^2\big]\\\leq&C_{s,1}^2\left(C_2^2+C_5\right)\|f\|_\infty^2\,\frac{s}{T}\big(\frac{1}{d}\sum_{j=1}^{d}\mathbb{E}\big[|\Delta\theta_j^0|^2\big]+\frac{1}{d}\sum_{j=1}^{d}\mathbb{E}\big[|\Delta\theta_j^t-\Delta\theta_j^0|^2\big]\big)\\&\to0,\qquad\text{as } d,T\to\infty,\end{aligned}$$

其中，第一个不等式由假设 1.2 得出。最后一个极限由命题 4.1以及定理 1.6中的初始条件得出。因此，我们已经证明 $N_{i,\ 3}^s\to0$ 依概率成立，记作 $N_{i,\ 3}^s=o$，并在下面的计算中将其标为 (1)。结合我们关于$N_{i,1}^s$、$N_{i,2}^s$、$N_{i,3}^s$的近似，对于任意 $i\in[d]$ 和 $s\in[0,\tau]$，我们可以将 (33) 的右端写为

$$\begin{aligned}&\mathrm{RHS}(33)\\=&-(i+1-dx)(N_{i,1}^s+N_{i,2}^s+N_{i,3}^s)-(dx-i)(N_{i+1,1}^s+N_{i+1,2}^s+N_{i+1,3}^s)\\=&-\eta dT\int_0^{\frac{\lfloor sT\rfloor}{T}}\int_0^1 f(u)((i+1-dx)A(\frac{i}{d},y)+(dx-i)A(\frac{i+1}{d},y))\Theta(u,y)\mathrm{d}y\mathrm{d}u\end{aligned}$$

$$
\begin{aligned}
&+\sigma^2\eta^2T\big((i+1-dx)\int_0^s f(u)\mathrm{d}\xi_1(u,\frac{i}{d})+(dx-i)\int_0^s f(u)\mathrm{d}\xi_1(u,\frac{i+1}{d})\big)\\
&+o(\eta dT)+o(\sigma^2\eta^2T)\\
=&-\eta dT\int_0^{\frac{\lfloor sT\rfloor}{T}}\int_0^1 f(u)A(x,y)\Theta(u,y)\mathrm{d}y\mathrm{d}u+\sigma^2\eta^2T\int_0^s f(u)\mathrm{d}\xi_1(u,x)\\
&+o(\eta dT)+o(\sigma^2\eta^2T),
\end{aligned}
\tag{36}
$$

其中，最后一个等式由以下事实得出： $A\in C^2([0,1])$，且对于任意 $x\in[\frac{i}{d},\frac{i+1}{d})$，~~$y\in[\frac{j}{d},\frac{j+1}{d})$~~，

$$
\begin{aligned}
&(i+1-dx)(j+1-dy)A(\frac{i}{d},\frac{j}{d})+(i+1-dx)(dy-j)A(\frac{i}{d},\frac{j+1}{d})\\
&+(dx-i)(j+1-dy)A(\frac{i+1}{d},\frac{j}{d})+(dx-i)(dy-j)A(\frac{i+1}{d},\frac{j+1}{d})\\
&=A(x,y)+o(1).
\end{aligned}
$$

最后，利用（34）和（36），对于任意 $s\in[0,\tau]$,

$$
\begin{aligned}
&-\int_0^{\frac{\lfloor sT\rfloor}{T}}\Theta(u,x)f'(u)\mathrm{d}u+f(s)\Theta(\frac{\lfloor sT\rfloor}{T},x)-f(0)\Theta(0,x)+o(1)\\
=&-\eta dT\int_0^{\frac{\lfloor sT\rfloor}{T}}\int_0^1 f(u)A(x,y)\Theta(u,y)\mathrm{d}y\mathrm{d}u+\sigma^2\eta^2T\int_0^s f(u)\mathrm{d}\xi_1(u,x)\\
&+o(\eta dT)+o(\sigma^2\eta^2T).
\end{aligned}
$$

因此，令 $d,\ T\to\infty$ 可得，对于任意有界光滑测试函数 f 以及任意$s\in[0,\tau]$, $x\in[0,1]$,

$$
\begin{aligned}
&-\int_0^s\Theta(u,x)f'(u)\mathrm{d}u+f(s)\Theta(s,x)-f(0)\Theta(0,x)\\
=&-\eta dT\big(\int_0^s\int_0^1 f(u)A(x,y)\Theta(u,y)\mathrm{d}y\mathrm{d}u+o(1)\big)\\
&+\sigma^2\eta^2T\big(\int_0^s f(u)\mathrm{d}\xi_1(u,x)+o(1)\big).
\end{aligned}
$$

■

证明 定理 1.6的证明。 证明可由定理 3.3的结果推出 .

如第 3.1节所述，中等噪声设定和高噪声设定分别可根据定理 3.3的推论 2 和 3，通过分部积分直接得到。对于 $C([0,\tau];C([0,1]))$ 中解的唯一性和解的存在性由定理 3.6的 (a) 部分推出。对于低噪声设定，由于 f 是光滑的，我们知道

$$
\int_0^{(\cdot)}\Theta(u,x)f'(u)\mathrm{d}u\in C^1([0,\tau])\quad\text{and}\quad\int_0^{(\cdot)}f(u)A(x,y)\Theta(u,y)\mathrm{d}u\mathrm{d}y\in C^1([0,\tau]).
$$

因此，根据（22），$f(\cdot)\Theta(\cdot,x)\in C^1([0,\tau])$，对任意 $x\in[0,1]$ 均成立，这意味着 $\Theta(\cdot,x)\in C^1([0,\tau])$，对任意 $x\in[0,1]$均成立。然后，我们可以对（22）的左侧应用分部积分。因此，Θ满足（8）。

由于 A 满足假设1.2，对任意 Θ_1，$\Theta_2 \in C([0,\tau];\ C([0,1]))$，我们有

$$\begin{aligned}
&\alpha \sup_{x\in[0,1]} \left| \int_0^1 A(x,y)\left(\Theta_1(s,y)-\Theta_2(s,y)\right) \mathrm{d}y \right| \\
\leq\ &\alpha \sup_{x,y\in[0,1]} |A(x,y)| \sup_{y\in[0,1]} |\Theta_1(s,y)-\Theta_2(s,y)| \\
\leq\ &\alpha C_2 \sup_{y\in[0,1]} |\Theta_1(s,y)-\Theta_2(s,y)| .
\end{aligned}$$

因此，(8) 的右端关于 $\Theta(s,\ \cdot)$ 对任意 $s\in[0,\tau]$ 都满足利普希茨连续性。根据 Picard-Lindel¨ 定理（例如参见 Arnold (1992年)），对于任意初始条件 $C([0,\tau];\ C([0,1])$ 8 式在 $\Theta(0,\ \cdot)=\Theta_0(\cdot)\in C([0,1]$ 上存在唯一解。

根据方程（8）解的唯一性和定理 3.3，可知 $\{\Theta^{d,T}\}_{d\geq 1,\ T>0}$ 的每个子列都有一个进一步的子列弱收敛到唯一的 $\Theta\in C([0,\tau];\ C([0,1]))$. 因此，$\{\Theta^{d,T}\}_{d\geq 1,\ T>0}$ 在 $d,\ T\to\infty$ 时弱收敛到 Θ。■

引理 4.3. 设 $N^{d,\ T}_{l}\in C([0,\tau];\ C([0,1]))$ 为在定理 *3.3* 的证明中针对 $l=1, 2, 3$ 定义的插值函数。在定理 *3.3* 的条件下，序列 $\{(\Theta^{d,T}, N^{d,\ T}_{1}, N^{d,\ T}_{2}, N^{d,\ T}_{3})\}_{d\geq 1,T>0}$ 在空间 $C([0,\tau];\ C([0,1]))^{\otimes 4}$ 中是紧的。

引理的证明 *4.3*。由于乘积空间结构，只需证明以下各项分别是紧的：$\Theta^{d,\ T}$，$N^{d,\ T}_{1}$，$N^{d,\ T}_{2}$，$N^{d,\ T}_{3}$ 在 $C([0,\tau];\ C([0,1])$ 中的紧性。$\{\Theta^{d,T}\}$ 的紧性已在定理 3.2中证明。本证明中，我们证明 $\{N^{d,T}_{l}\}$ 在 $l=1$、2、3 下的紧性。

的紧性 $N^{d,\ T}_{1}$。首先，我们有 $N^{d,\ T}_{1}(0,\ 0)=-\eta f(0)\sum_{j=1}^{d}A(0,\ \frac{j}{d})\Theta^{d,\ T}(0,\ \frac{j}{d})$。在定理 3.3的假设下，我们有

$$\begin{aligned}
\mathbb{P}(|N_1^{d,T}(0,0)|>N) &\leq N^{-2}\eta^2 f(0)^2\Big(\sum_{j=1}^{d}A(0,\frac{j}{d})^2\Big)\Big(\sum_{j=1}^{d}\mathbb{E}[|\Delta\theta_j^0|^2]\Big) \\
&\leq C_2^2\,\|f\|_\infty^2\,N^{-2}\eta^2 d^2 R^2.
\end{aligned}$$

由于 $\eta dT\leq C_{s,\ 1}$ 且一致常数为 $C_{s,\ 1}$，$\mathbb{P}(|N^{d,\ T}_{1}(0,\ 0)|>N)\lesssim N^{-2}\to 0$ 当 $N\to\infty$ 时。因此，$\{N_1^{d,T}(0,\ 0)\}_{d\geq 1,\ T>0}$ 在概率空间中是紧的。接下来我们注意到，对于任意 $0\leq t_1<t_2\leq\lfloor sT\rfloor-1$ 和 $i_1,\ i_2\in[d]$,

$$\begin{aligned}
N_1^{d,T}(\frac{t_2}{T},\frac{i_2}{d})-N_1^{d,T}(\frac{t_1}{T},\frac{i_1}{d}) &= \eta\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d}f(\frac{t}{T})A(\frac{i_2}{d},\frac{j}{d})\overline{\Theta}^{d,T}(\frac{t}{T},\frac{j}{d}) \\
&\quad+\eta\sum_{t=0}^{t_1-1}\sum_{j=1}^{d}f(\frac{t}{T})\big(A(\frac{i_2}{d},\frac{j}{d})-A(\frac{i_1}{d},\frac{j}{d})\big)\overline{\Theta}^{d,T}(\frac{t}{T},\frac{j}{d})
\end{aligned}$$

因此，我们有

$$\begin{aligned}
&\mathbb{E}[|N_1^{d,T}(\frac{t_2}{T},\frac{i_2}{d})-N_1^{d,T}(\frac{t_1}{T},\frac{i_1}{d})|^4] \\
\lesssim\ &\eta^4\mathbb{E}\Big[\sum_{r_1,r_2,r_3,r_4=t_1}^{t_2-1}\ \sum_{j_1,j_2,j_3,j_4=1}^{d}\ \prod_{k=1}^{4}|A(\frac{i_2}{d},\frac{j_k}{d})|\prod_{k=1}^{4}|f(\frac{r_k}{T})\Delta\theta_{j_k}^{r_k}|\Big]
\end{aligned}$$

$$
\begin{aligned}
&+\eta^4\mathbb{E}\big[\sum_{r_1,r_2,r_3,r_4=0}^{t_1-1}\sum_{j_1,j_2,j_3,j_4=1}^{d}\prod_{k=1}^{4}|A(\frac{i_2}{d},\frac{j_k}{d})-A(\frac{i_1}{d},\frac{j_k}{d})|\prod_{k=1}^{4}|f(\frac{r_k}{T})\Delta\theta_{j_k}^{r_k}|\big]\\
&\le C_2^4\|f\|_\infty^4\eta^4\mathbb{E}\big[\sum_{r_1,r_2,r_3,r_4=t_1}^{t_2-1}\sum_{j_1,j_2,j_3,j_4=1}^{d}\prod_{k=1}^{4}|\Delta\theta_{j_k}^{r_k}|\big]\\
&+C_2^4\|f\|_\infty^4\eta^4\big(\frac{i_2-i_1}{d}\big)^4\mathbb{E}\big[\sum_{r_1,r_2,r_3,r_4=0}^{t_1-1}\sum_{j_1,j_2,j_3,j_4=1}^{d}\prod_{k=1}^{4}|\Delta\theta_{j_k}^{r_k}|\big]\\
&\le C_2^4\|f\|_\infty^4\eta^4d^4(t_2-t_1)^3\sum_{t=t_1}^{t_2-1}m_t^d\\
&+C_2^4C_3^4\|f\|_\infty^4\eta^4d^4t_1^3\big(\frac{i_2-i_1}{d}\big)^4\sum_{t=0}^{t_1-1}m_t^d\\
&\lesssim(\eta dT)^4\big(\frac{t_2-t_1}{T}\big)^4+(\eta dT)^4\big(\frac{i_2-i_1}{d}\big)^4,
\end{aligned}
$$

其中第三个不等式由假设 1.2 得出，最后一个不等式则由命题 4.1证明中导出的 m_t^d 上界得出。由于 $N^{d,\cdot,1,T}$ 关于两个变量均为分段线性函数，且对于所有 $d\ge1$，$T>0$，都有 $\eta dT\le C_{s,1}$，因此对于所有 $0\le s_1<s_2\le\tau$ 和 $0\le x_1<x_2\le1$，我们有

$$
\mathbb{E}\big[\big|N_1^{d,T}(s_1,x_1)-N_1^{d,T}(s_2,x_2)\big|^4\big]\lesssim(s_2-s_1)^4+(x_1-x_2)^4.
$$

$N^{d,\cdot,1,T}$在 $C([0,\tau];C([0,1]))$ 中的紧性由柯尔莫哥洛夫紧性判据得出。

.....的紧性 $N^{d,\cdot,2,T}$。首先，我们有 $N^{d,\cdot,2,T}(0,0)=-\eta f(0)W(0,0)\varepsilon^0$。根据定理3.3中的假设，我们有

$$
\mathbb{P}(|N_2^{d,T}(0,0)|>N)\le N^{-2}\eta^2f(0)^2\sigma^2A(0,0)\le C_2\|f\|_\infty^2N^{-2}\eta^2\sigma^2.
$$

由于 $\sigma\eta T^{\frac12}\le C_{s,1}$对于一致常数$C_{s,1}$，$\mathbb{P}(|N^{d,\cdot,2,T}(0,0)|>N)\lesssim N^{-2}\to0$当 $N\to\infty$时。因此，$\{N_2^{d,T}(0,0)\}_{d\ge1,T>0}$在概率空间中是紧的。接下来，我们注意到，对于任意 $0\le t_1<t_2\le\lfloor sT\rfloor-1$和 $i_1,i_2\in[d]$，

$$
\begin{aligned}
N_2^{d,T}(\frac{t_2}{T},\frac{i_2}{d})-N_2^{d,T}(\frac{t_1}{T},\frac{i_1}{d})&=\eta\sum_{t=t_1}^{t_2-1}f(\frac{t}{T})W(\frac{t}{T},\frac{i_2}{d})\varepsilon^t\\
&+\eta\sum_{t=0}^{t_1-1}f(\frac{t}{T})\big(W(\frac{t}{T},\frac{i_2}{d})-W(\frac{t}{T},\frac{i_1}{d})\big)\varepsilon^t
\end{aligned}
$$

由命题 $\mathbb{E}[S_{i,3}]$和 $\mathbb{E}[S_{i,j,3}]$ 的证明中对4.1 以及命题4.2 进行界定的论证可知

$$
\begin{aligned}
\mathbb{E}[|N_2^{d,T}(\frac{t_2}{T},\frac{i_2}{d})-N_2^{d,T}(\frac{t_1}{T},\frac{i_1}{d})|^4]&\lesssim C_2^2\eta^4\|f\|_\infty^4\sigma^4(t_2-t_1)^2+C_5\eta^4\|f\|_\infty^4\sigma^4(t_2-t_1)\\
&+C_7^4\eta^4\|f\|_\infty^4\sigma^4t_1^4\big(\frac{i_2-i_1}{d}\big)^4\\
&\lesssim(\sigma\eta T^{\frac12})^4\big(\frac{t_2-t_1}{T}\big)^2+(\sigma\eta T^{\frac12})^4\big(\frac{i_2-i_1}{d}\big)^4.
\end{aligned}
$$

由于 $N^{d,T}_{\cdot,2}$关于两个变量均为分段线性函数，且 $\sigma\eta T^{\frac{1}{2}} \le C_{s,1}$对所有 $d \ge 1$，$T > 0$以及所有 $0 \le s_1 < s_2 \le \tau$和 $0 \le x_1 < x_2 \le 1$均成立，因此

$$\mathbb{E}\big[\big|N_2^{d,T}(s_1,x_1) - N_2^{d,T}(s_2,x_2)\big|^4\big] \lesssim (s_2-s_1)^2 + (x_1-x_2)^4.$$

紧性　　由柯尔莫哥洛夫紧性判据可得，$N^{d,T}_{\cdot,2}$在 $C([0,\tau]; C([0,1])$ 中

.....的紧性 $N^{d,T}_{\cdot,3}$。首先，我们有 $N^{d,T}_{\cdot,3}(0,0) = -\eta f(0)\sum_{j=1}^{d}\big(W(0,0)W(0,\frac{j}{d}) - A(0,\frac{j}{d})\big)\Theta^{d,T}(0,\frac{j}{d})$. 在定理3.3的假设下，我们有

$$\begin{aligned}\mathbb{P}(|N_3^{d,T}(0,0)| > N) &\le N^{-2}\eta^2 f(0)^2 \sum_{j=1}^{d}\big(B(0,\frac{j}{d},0,\frac{j}{d}) - A(0,\frac{j}{d})^2\big)\sum_{j=1}^{d}\mathbb{E}[|\Delta\theta_j^0|^2]\\ &\le (C_2^2 + C_5)\,\|f\|_\infty^2\, N^{-2}\eta^2 d^2 R^2.\end{aligned}$$

由于 $\eta dT \le C_{s,1}$其中 $C_{s,1}$为一致常数，$\mathbb{P}(|N^{d,T}_{\cdot,3}(0,0)| > N) \lesssim N^{-2} \to 0$当 $N \to \infty$时。因此，$\{N_3^{d,T}(0,0)\}_{d\ge1,T>0}$在该概率空间中是紧的。接下来，我们注意到，对任意 $0 \le t_1 < t_2 \le \lfloor sT\rfloor - 1$和 $i_1, i_2 \in [d]$,

$$\begin{aligned}&N_3^{d,T}(\frac{t_2}{T},\frac{i_2}{d}) - N_3^{d,T}(\frac{t_1}{T},\frac{i_1}{d})\\ &= \eta\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d} f(\frac{t}{T})\big(W(\frac{t}{T},\frac{i_2}{d})W(\frac{t}{T},\frac{j}{d}) - A(\frac{i_2}{d},\frac{j}{d})\big)\Theta^{d,T}(\frac{t}{T},\frac{j}{d})\\ &+ \eta\sum_{t=0}^{t_1-1}\sum_{j=1}^{d} f(\frac{t}{T})\Big(W(\frac{t}{T},\frac{i_2}{d})W(\frac{t}{T},\frac{j}{d}) - W(\frac{t}{T},\frac{i_1}{d})W(\frac{t}{T},\frac{j}{d})\\ &\qquad - A(\frac{i_2}{d},\frac{j}{d}) + A(\frac{i_1}{d},\frac{j}{d})\Big)\Theta^{d,T}(\frac{t}{T},\frac{j}{d}).\end{aligned}$$

根据命题 $\mathbb{E}[S_{i,2}]$和 $\mathbb{E}[S_{i,j,2}]$的证明中对 4.1 和命题4.2 的界估计论证可得

$$\begin{aligned}\mathbb{E}[|N_3^{d,T}(\frac{t_2}{T},\frac{i_2}{d}) - N_3^{d,T}(\frac{t_1}{T},\frac{i_1}{d})|^4] &\lesssim \eta^4\|f\|_\infty^4 d^4(t_2-t_1)^2 + \eta^4\|f\|_\infty^4 d^4 t_1^2(\frac{i_2-i_1}{d})^4\\ &\lesssim (\eta dT)^4(\frac{t_2-t_1}{T})^2 + (\eta dT)^4(\frac{i_2-i_1}{d})^4.\end{aligned}$$

由于 $N^{d,T}_{\cdot,3}$关于两个变量都是分段线性的，并且 $\eta dT \le C_{s,1}$对所有 $d \ge 1$，$T > 0$以及所有 $0 \le s_1 < s_2 \le \tau$和 $0 \le x_1 < x_2 \le 1$均成立，因此我们有

$$\mathbb{E}\big[\big|N_3^{d,T}(s_1,x_1) - N_3^{d,T}(s_2,x_2)\big|^4\big] \lesssim (s_2-s_1)^2 + (x_1-x_2)^4.$$

$N^{d,T}_{\cdot,3}$在 $C([0,\tau]; C([0,1])$ 中的紧性可由柯尔莫哥洛夫紧性判据推出。 ■

5 涨落的证明

在介绍定理 3.5 的证明之前，我们先考察涨落的迭代公式，并证明涨落的紧性。根据 $\Theta^{d,T}$的定义，在定理 1.6 的低噪声设定下，对于任意 $0 \le t \le \lfloor \tau T\rfloor - 1 := N-1$，$i \in [d]$，有

$$
\begin{aligned}
&U^{d,T}(\frac{t+1}{T},\frac{i}{d})-U^{d,T}(\frac{t}{T},\frac{i}{d})\\
=&-\eta\sum_{j=1}^{d}W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})U^{d,T}(\frac{t}{T},\frac{j}{d})\\
&-\eta\gamma\sum_{j=1}^{d}\big(W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})-\mathbb{E}[W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})]\big)\Theta(\frac{t}{T},\frac{j}{d})\\
&-\gamma\big(\Theta(\frac{t+1}{T},\frac{i}{d})-\Theta(\frac{t}{T},\frac{i}{d})+\eta\sum_{j=1}^{d}A(\frac{i}{d},\frac{j}{d})\Theta(\frac{t}{T},\frac{j}{d})\big)\\
&+\eta\gamma W(\frac{t}{T},\frac{i}{d})\varepsilon^{t}.
\end{aligned}
\tag{37}
$$

5.1 $\{U^{d,T}\}_{d\geq 1,\ T>0}$的紧性

命题 5.1。 设假设 *1.1* 和 *1.2* 成立，并进一步假设定理*1.6* 和定理 *1.10* 中的初始条件均满足。若存在一致正常数$C_{s,\ 2}$，使得沿极限序列对所有 $d\geq 1$、$T>0$均有 $\max(\gamma T^{-\frac{1}{2}}$、$\gamma d^{-1}$、$\gamma\sigma d^{-1}T^{-\frac{1}{2}})\leq C_{s,\ 2}$，则存在一致常数 $C>0$，使得对于任意 $d\geq 1$、$T>0$，任意 $i\in[d]$ 和任意$0\leq t_1<t_2\leq\lfloor\tau T\rfloor$,

$$
\mathbb{E}\big[\big|U^{d,T}(\frac{t_2}{T},\frac{i}{d})-U^{d,T}(\frac{t_1}{T},\frac{i}{d})\big|^4\big]\leq C\left(\frac{t_2-t_1}{T}\right)^2.
\tag{38}
$$

证明 命题 *5.1*。 由（37），可知对任意 $0\leq t_1<t_2\leq N:=\lfloor\tau T\rfloor$,

$$
\begin{aligned}
&U^{d,T}(\frac{t_2}{T},\frac{i}{d})-U^{d,T}(\frac{t_1}{T},\frac{i}{d})\\
=&\underbrace{-\eta\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d}A(\frac{i}{d},\frac{j}{d})U^{d,T}(\frac{t}{T},\frac{j}{d})}_{M_{i,1}}\\
&\underbrace{-\eta\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d}\big(W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})-\mathbb{E}[W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})]\big)U^{d,T}(\frac{t}{T},\frac{j}{d})}_{M_{i,2}}\\
&\underbrace{-\eta\gamma\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d}\big(W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})-\mathbb{E}[W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})]\big)\Theta(\frac{t}{T},\frac{j}{d})}_{M_{i,3}}\\
&\underbrace{-\gamma\big(\Theta(\frac{t_2}{T},\frac{i}{d})-\Theta(\frac{t_1}{T},\frac{i}{d})+\eta\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d}A(\frac{i}{d},\frac{j}{d})\Theta(\frac{t}{T},\frac{j}{d})\big)}_{M_{i,4}}+\underbrace{\eta\gamma\sum_{t=t_1}^{t_2-1}W(\frac{t}{T},\frac{i}{d})\varepsilon^{t}}_{M_{i,5}}.
\end{aligned}
\tag{39}
$$

因此，我们有

$$\mathbb{E}\big[\big|U^{d,T}(\frac{t_2}{T},\frac{i}{d})-U^{d,T}(\frac{t_1}{T},\frac{i}{d})\big|^4\big]\le 125\sum_{j=1}^{5}\mathbb{E}[M_{i,j}^4]. \tag{40}$$

接下来，我们将逐项估计右侧项的期望。许多项可以利用命题 4.1的证明进行估计。定义 n_t^d:
$=\frac{1}{d}\sum_{i=1}^{d}\mathbb{E}\big[\big|U^{d,T}(\frac{t}{T},\frac{i}{d})\big|^4\big]$。 $\mathbb{E}[M_{i,1}^4]$可以用类似于 $\mathbb{E}[S_{i,1}]$ 命题 4.1证明中的方法进行估计，并且有

$$\begin{aligned}\mathbb{E}\big[M_{i,1}^4\big]&=\eta^4\mathbb{E}\big[\sum_{r_1,r_2,r_3,r_4=t_1}^{t_2-1}\sum_{j_1,j_2,j_3,j_4=1}^{d}\prod_{k=1}^{4}A(\frac{i}{d},\frac{j_k}{d})\prod_{k=1}^{4}U^{d,T}(\frac{r_k}{T},\frac{j_k}{d})\big]\\&\le C_2^4\eta^4d^4(t_2-t_1)^3\sum_{t=t_1}^{t_2-1}n_t^d.\end{aligned}$$

$\mathbb{E}[M_{i,2}^4]$ 可以用类似于 $\mathbb{E}[S_{i,2}]$ 命题 4.1证明中的方法进行估计，并且有

$$\mathbb{E}\big[M_{i,2}^4\big]\le 3\left(C_8'+C_5'^2\right)\eta^4d^4(t_2-t_1)\sum_{t=t_1}^{t_2-1}n_t^d.$$

类似于对 $\mathbb{E}[M_{i,2}^4]$的估计，我们有

$$\begin{aligned}\mathbb{E}\big[M_{i,3}^4\big]&\le 3\big(C_8'+C_5'^2\big)(t_2-t_1)\eta^4\gamma^4\mathbb{E}\big[\sum_{t=t_1}^{t_2-1}\sum_{j_1,j_2,j_3,j_4=1}^{d}\prod_{k=1}^{4}\Theta(\frac{t}{T},\frac{j_k}{d})\big]\\&\le 3\left(C_8'+C_5'^2\right)\|\Theta\|_\infty^4\gamma^4\eta^4d^4(t_2-t_1)^2\\&=O(\gamma^4\eta^4d^4(t_2-t_1)^2),\end{aligned}$$

其中，第二个不等式由以下事实得出:

$$\|\Theta\|_\infty:=\sup_{s\in[0,\tau],x\in[0,1]}|\Theta(s,x)|<\infty.$$

接下来，由于 A，Θ 在所有变量上都是 C^1，因此有

$$\begin{aligned}\mathbb{E}\left[M_{i,4}^2\right]&=\gamma^4\eta^4d^4T^4\big(-\int_{\frac{t_1}{T}}^{\frac{t_2}{T}}\int_0^1A(\frac{i}{d},y)\Theta(s,y)\mathrm{d}y\mathrm{d}s+\frac{1}{dT}\sum_{t=t_1}^{t_2-1}\sum_{j=1}^{d}A(\frac{i}{d},\frac{j}{d})\Theta(\frac{t}{T},\frac{j}{d})\big)^4\\&=O(\gamma^4\eta^4(t_2-t_1)^4)+O(\gamma^4\eta^4d^4(t_2-t_1)^4T^{-4}).\end{aligned}$$

$\mathbb{E}[M_{i,5}^4]$ 可以用类似于 $\mathbb{E}[S_{i,3}]$ 命题 4.1证明中的方法进行估计，并且有

$$\mathbb{E}\left[M_{i,5}^4\right]\le(C_5C_1+C_2^2)\sigma^4\eta^4\gamma^4(t_2-t_1)^2=O(\gamma^4\eta^4\sigma^4(t_2-t_1)^2).$$

结合所有估计并取 $t_2=t$，$t_1=0$ 在（40）中，

$$n_t^d=\frac{1}{d}\sum_{i=1}^{d}\mathbb{E}\big[\big|U^{d,T}(\frac{t}{T},\frac{i}{d})\big|^4\big]\le 8n_0^d+\frac{8}{d}\sum_{i=1}^{d}\mathbb{E}\big[\big|U^{d,T}(\frac{t}{T},\frac{i}{d})-U^{d,T}(0,\frac{i}{d})\big|^4\big]$$

$$\leq 8n_0^d + 2000\big(C_2^4 + 3(C_8 + C_5^2)\big)\eta^4 d^4 t^3 \sum_{k=0}^{t-1} n_k^d + C_{d,T}(t),$$

其中 $C_{d,\ T}(t) := O(\gamma^4\eta^4 d^4 t^2 + \gamma^4\eta^4 t^4 + \gamma^4\eta^4 d^4 t^4 T^{-4} + \gamma^4\eta^4\sigma^4 t^2)$。根据离散 Gronwall 不等式，我们有

$$n_t^d \leq C_{d,T}(t) + \tilde{C}\eta^4 d^4 t^3 \sum_{k=0}^{t-1} C_{d,T}(k) \exp\big(\tilde{C}\eta^4 d^4 t^3 (t-k-1)\big),$$

其中 $\tilde{C} = 2000(C_2^4 + 3(C_8 + C_5^2))$. 由于 $\eta dT \leq C_{s,\ 1}$，$2000\big(C_2^4 + 3(C_8 + C_5^2)\big)\eta^4 d^4 t^3 (t-k-1) \leq C_\tau C_{s,\ 1}^4 \tau^4$ 对于任意 $0 \leq k < t$、$t \leq N$和 $C_{\tau,\ 1}$都是与 d 无关的常数，T，σ。由于 $C_{d,\ T}(k)$ 随 k 增大而增大，因此存在一个与 d 无关的常数 $C_{\tau,\ 2}$，T，σ 使得

$$\begin{aligned} n_t^d &\leq C_{d,T}(t) + 2000\big(C_2^4 + 3(C_8 + C_5^2)\big) \exp\left(C_{\tau,1} C_{s,1}^4 \sigma^{-4}\tau^4\right) \eta^4 d^4 t^3 \sum_{k=0}^{t-1} C_{d,T}(k) \\ &\leq C_{\tau,2} C_{d,T}(t). \end{aligned} \tag{41}$$

将（41）代入（40），并取期望。于是得到

$$\begin{aligned} &\mathbb{E}\big[\big|U^{d,T}(\frac{t_2}{T}, \frac{i}{d}) - U^{d,T}(\frac{t_1}{T}, \frac{i}{d})\big|^4\big] \\ \leq& 125\big(C_2^4 + 3(C_8' + C_5'^2)\big)\eta^4 d^4 (t_2 - t_1)^3 \sum_{t=t_1}^{t_2-1} n_t^d + C_{d,T}(t_2 - t_1) \\ \leq& 125 C_{\tau,2}\big(C_2^4 + 3(C_8' + C_5'^2)\big)\eta^4 d^4 (t_2 - t_1)^3 \sum_{t=t_1}^{t_2-1} C_{d,T}(t) + C_{d,T}(t_2 - t_1). \end{aligned}$$

注意到，由于 $\max(\eta dT,\ \eta\sigma T^{\frac{1}{2}}) \leq C_{s,\ 1}$，我们有

$$C_{d,T}(t) = \big(\frac{t}{T}\big)^2 O(\gamma^4 T^{-2} + \gamma^4 d^{-4} + \gamma^4\sigma^4 d^{-4} T^{-2}).$$

此外，注意到

$$\max(\gamma T^{-\frac{1}{2}}, \gamma d^{-1}, \gamma\sigma d^{-1} T^{-\frac{1}{2}}) \leq C_{s,2} \quad \text{and} \quad C_{d,T}(t) \leq C_{\tau,3} C_{s,2}^4 \big(\frac{t}{T}\big)^2$$

对于某个与 d，T，σ，γ 无关的正常数 $C_{\tau,\ 3}$。因此，

$$\begin{aligned} &\mathbb{E}\big[\big|U^{d,T}(\frac{t_2}{T}, \frac{i}{d}) - U^{d,T}(\frac{t_1}{T}, \frac{i}{d})\big|^4\big] \\ \leq& 125 C_{\tau,2}\big(C_2^4 + 3(C_8' + C_5'^2)\big)\eta^4 d^4 (t_2 - t_1)^4 C_{\tau,3} C_{s,2}^4 \tau^2 + C_{\tau,3}\big(\frac{t_2 - t_1}{T}\big)^2 \\ \leq& 125 C_{\tau,2}\big(C_2^4 + 3(C_8' + C_5'^2)\big) C_{s,1}^4 C_{\tau,3} C_{s,2}^4 \tau^2 \big(\frac{t_2 - t_1}{T}\big)^4 + C_{\tau,3}\big(\frac{t_2 - t_1}{T}\big)^2. \end{aligned}$$

最后的（43）可由 $\big(\frac{t_2 - t_1}{T}\big)^2 \leq \tau^2$得证。 ■

命题 5.2。 设假设 *1.1* 和 *1.2*成立，且进一步假设定理 *1.6* 中的初始条件满足。对于任意 d， $T>0$，存在一个与 $C>0$ 无关的常数 d， T使得对于任意 $d\geq 1$， $T>0$，任意 i， $j\in[d]$ 以及任意 $0\leq t_1<t_2\leq\lfloor\tau T\rfloor$，

$$\mathbb{E}\big[\big|\big(U^{d,T}(\frac{t_2}{T},\frac{i}{d})-U^{d,T}(\frac{t_2}{T},\frac{j}{d})\big)-\big(U^{d,T}(\frac{t_1}{T},\frac{i}{d})-U^{d,T}(\frac{t_1}{T},\frac{j}{d})\big)\big|^4\big]\leq C\big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^2. \tag{42}$$

此外，若定理 *1.10* 中的初始条件成立，则对任意 i、 $j\in[d]$ 和任意 $0\leq t\leq\lfloor\tau T\rfloor$，

$$\mathbb{E}\big[\big|U^{d,T}(\frac{t}{T},\frac{i}{d})-U^{d,T}(\frac{t}{T},\frac{j}{d})\big|^4\big]\leq 8\left(C\tau^2+M^4\right)\big(\frac{i-j}{d}\big)^4, \tag{43}$$

其中 M 是定理 *1.10*中初始条件所对应的一致常数。

命题 *5.2* 的证明。由（11）可知，对任意 $0\leq t_1<t_2\leq N:=\lfloor\tau T\rfloor$ 以及任意 i， $j\in[d]$，

$$\begin{aligned}
&\big(U^{d,T}(\tfrac{t_2}{T},\tfrac{i}{d})-U^{d,T}(\tfrac{t_2}{T},\tfrac{j}{d})\big)-\big(U^{d,T}(\tfrac{t_1}{T},\tfrac{i}{d})-U^{d,T}(\tfrac{t_1}{T},\tfrac{j}{d})\big)\\
=&\underbrace{-\eta\sum_{t=t_1}^{t_2-1}\sum_{l=1}^{d}\big(A(\tfrac{i}{d},\tfrac{l}{d})-A(\tfrac{j}{d},\tfrac{l}{d})\big)U^{d,T}(\tfrac{t}{T},\tfrac{l}{d})}_{M_{i,j,1}}\\
&\underbrace{-\eta\sum_{t=t_1}^{t_2-1}\sum_{l=1}^{d}\left(x_i^tx_l^t-\mathbb{E}\left[x_i^tx_l^t\right]-x_j^tx_l^t+\mathbb{E}\left[x_j^tx_l^t\right]\right)U^{d,T}(\tfrac{t}{T},\tfrac{l}{d})}_{M_{i,j,2}}\\
&\underbrace{+\alpha\gamma\int_{\frac{t_1}{T}}^{\frac{t_2}{T}}\int_0^1\big(A(\tfrac{i}{d},y)-A(\tfrac{j}{d},y)\big)\Theta(s,y)dyds-\eta\gamma\sum_{l=1}^{d}\big(A(\tfrac{i}{d},\tfrac{l}{d})-A(\tfrac{j}{d},\tfrac{l}{d})\big)\Theta(\tfrac{t}{T},\tfrac{l}{d})}_{M_{i,j,3}}\\
&\underbrace{-\eta\gamma\sum_{t=t_1}^{t_2-1}\sum_{l=1}^{d}\left(x_i^tx_l^t-\mathbb{E}\left[x_i^tx_l^t\right]-x_j^tx_l^t+\mathbb{E}\left[x_j^tx_l^t\right]\right)\Theta(\tfrac{t}{T},\tfrac{l}{d})}_{M_{i,j,4}}+\underbrace{\eta\gamma\sum_{t=t_1}^{t_2-1}\left(x_i^t-x_j^t\right)\varepsilon^t}_{M_{i,j,5}}.
\end{aligned}$$

因此，若定义 $\Delta^{d,T}(t,i,j):=U^{d,T}(\frac{t}{T},\frac{i}{d})-U^{d,T}(\frac{t}{T},\frac{j}{d})$，则有

$$\mathbb{E}\big[\big|\Delta^{d,T}(t_2,i,j)-\Delta^{d,T}(t_1,i,j)\big|^4\big]\leq 125\sum_{l=1}^{5}\mathbb{E}\big[M_{i,j,l}^4\big]. \tag{44}$$

接下来，我们逐项估计（44）右端的各项；根据命题 4.2 的证明，大多数项都有界。 $\mathbb{E}[M_{i,j,1}^4]$ 可用与命题 $\mathbb{E}[S_{i,j,1}]$ 4.2 证明中相似的方法作上界估计，且我们有

$$\mathbb{E}\big[M_{i,j,1}^4\big]\leq C_3^4\big(\frac{i-j}{d}\big)^4(t_2-t_1)^3\eta^4d^4\sum_{t=t_1}^{t_2-1}n_t^d=\big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^2O(\eta^4d^4T^3\sum_{t=t_1}^{t_2-1}n_t^d).$$

$\mathbb{E}[M_{i,j,2}^4]$ 可以像 $\mathbb{E}[S_{i,j,2}]$ 命题4.2的证明中那样进行上界估计，并且我们有

$$\mathbb{E}\big[M_{i,j,2}^4\big]\leq\big(C_9'^4+C_6'^4\big)\big(\frac{i-j}{d}\big)^4(t_2-t_1)\eta^4d^4\sum_{t=t_1}^{t_2-1}n_t^d$$

$$= \big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^2 O(\eta^4 d^4 T^2 \sum_{t=t_1}^{t_2-1} n_t^d).$$

由于 Θ 关于两个变量都是 C^1，且对于任意 $x \in [0,1]$，$\partial_1 A(x,\ \cdot) \in C^1([0,1])$，因此 $\mathbb{E}[M_{i,j,3}^4]$ 可以估计为

$$\begin{aligned}
&\mathbb{E}\big[M_{i,j,3}^4\big] \\
&= \alpha^4\gamma^4\Big(\int_{\frac{i}{d}}^{\frac{j}{d}}\int_{\frac{t_1}{T}}^{\frac{t_2}{T}}\int_0^1 \partial_1 A(z,y)\Theta(s,y)\mathrm{d}y\mathrm{d}s\mathrm{d}z - \int_{\frac{i}{d}}^{\frac{j}{d}} \frac{1}{dT}\sum_{l=1}^{d}\partial_1 A(z,\frac{l}{d})\Theta(\frac{t}{T},\frac{l}{d})\mathrm{d}z\Big)^4 \\
&= O\bigg(\gamma^4\big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^4\frac{1}{T^4}\bigg) + O\bigg(\gamma^4\big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^4\frac{1}{d^4}\bigg) \\
&= \big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^4 O\Big(\gamma^4 T^{-4} + \gamma^4 d^{-4}\Big).
\end{aligned}$$

接下来，$\mathbb{E}[M_{i,j,4}^4]$ 可以用与估计 $\mathbb{E}[M_{i,j,2}^4]$ 类似的方法进行估计：

$$\begin{aligned}
\mathbb{E}\big[M_{i,j,4}^4\big] &\le C_9'^4\eta^4\gamma^4\big(\frac{i-j}{d}\big)^4\sum_{t=t_1}^{t_2-1}\sum_{l_1,l_2,l_3,l_4=1}^{d}\mathbb{E}\big[\big|\prod_{k=1}^{4}\Theta(\frac{t}{T},\frac{l_k}{d})\big|\big] \\
&\quad + C_6'^4\eta^4\gamma^4\big(\frac{i-j}{d}\big)^4\sum_{t',t=t_1}^{t_2-1}\sum_{l_1,l_2,l_3,l_4=1}^{d}\mathbb{E}\big[\big|\prod_{k=1}^{2}\Theta(\frac{t}{T},\frac{l_k}{d})\prod_{k=3}^{4}\Theta(\frac{t'}{T},\frac{l_k}{d})\big|\big] \\
&\le \big(C_9'^4 + C_6'^4\big)\|\Theta\|_\infty^4\big(\frac{i-j}{d}\big)^4\eta^4\gamma^4 d^4(t_2-t_1)^2 \\
&= \big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^2 O(\eta^4\gamma^4 d^4 T^2).
\end{aligned}$$

最后，$\mathbb{E}[M_{i,j,5}^4]$ 可以 用与 $\mathbb{E}[S_{i,j,3}]$ 命题的证明4.2中类似的方法进行上界估计，我们有 e

$$\mathbb{E}\big[M_{i,j,5}^4\big] \le \big(C_1 C_7^4 + C_4^4\big)\big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^2\eta^4\gamma^4\sigma^4 T^2.$$

在命题5.1的证明中，我们已证明$n_t^d \le C_{\tau,\ 2}C_{d,\ T}(t) \le C_{\tau,\ 2}C_{\tau,\ 3}C_{s,\ 2}\big(\frac{t}{T}\big)^2 = O(1)$。因此，将此估计应用于（44），并结合对 $\mathbb{E}[M_{i,j,\cdot}^4]$ 的估计，可得

$$\begin{aligned}
&\mathbb{E}\big[\big|\Delta^{d,T}(t_2,i,j) - \Delta^{d,T}(t_1,i,j)\big|^4\big] \\
&\le \big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^2 O\bigg(\eta^4 d^4 T^3\sum_{t=t_1}^{t_2-1} n_t^d + \gamma^4 T^{-4} + \gamma^4 d^{-4} + \eta^4\gamma^4 d^4 T^2 + \eta^4\gamma^4\sigma^4 T^2\bigg) \\
&= \big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^2 O\big(\eta^4 d^4 T^4 + \gamma^4 T^{-4} + \gamma^4 d^{-4} + \eta^4\gamma^4 d^4 T^2 + \eta^4\gamma^4\sigma^4 T^2\big).
\end{aligned}$$

由于 $\max(\eta d T,\ \eta\sigma T^{\frac{1}{2}}) \le C_{s,\ 1}$ 且 $\max(\gamma T^{-\frac{1}{2}},\ \gamma d^{-1},\ \gamma\sigma d^{-1}T^{-\frac{1}{2}}) \le C_{s,\ 2}$，

$$O\big(\eta^4 d^4 T^4 + \gamma^4 T^{-4} + \gamma^4 d^{-4} + \eta^4\gamma^4 d^4 T^2 + \eta^4\gamma^4\sigma^4 T^2\big) = O(1),$$

因此存在一致常数 C，使得

$$\mathbb{E}\big[\big|\Delta^{d,T}(t_2,i,j) - \Delta^{d,T}(t_1,i,j)\big|^4\big] \le C\big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^2.$$

取 $t_2 = t$ 和 $t_1 = 0$，得到

$$\begin{aligned}
&\mathbb{E}\big[\big|U^{d,T}(\frac{t}{T},\frac{i}{d}) - U^{d,T}(\frac{t}{T},\frac{j}{d})\big|^4\big] \\
&\le 8\mathbb{E}\big[\big|U^{d,T}(0,\frac{i}{d}) - U^{d,T}(0,\frac{j}{d})\big|^4\big] + 8C\big(\frac{i-j}{d}\big)^4\big(\frac{t_2-t_1}{T}\big)^2 \\
&\le 8\big(C\tau^2 + M^4\big)\big(\frac{i-j}{d}\big)^4,
\end{aligned}$$

最后一个不等式由定理 1.10 中的初始条件得到。 ■

定理 *3.4* 的证明。 当 $\eta = \frac{\alpha}{dT}$ 且 $\sigma = o(dT^{\frac{1}{2}})$ 时，存在一个一致常数 $C_{s,\,1}$ 使得对于所有 $d \ge 1$，$T > 0$ 沿着极限序列，$\max(\eta dT,\ \eta\sigma T^{\frac{1}{2}}) \le C_{s,\,1}$。因此，根据命题 5.1 和命题 5.2，定理 3.4 的证明与定理 3.2 的证明完全相同。因此，我们略去细节。 ■

5.2 极限识别

定理 *3.5* 的证明。 对于任意 $0 \le t \le \lfloor \tau T \rfloor - 1 := N - 1$，$i \in [d-1]$，根据 (3) 和 (11)，对所有 $x \in [\frac{i}{d}, \frac{i+1}{d}]$，

$$\begin{aligned}
&U^{d,T}(\frac{t+1}{T}, x) - U^{d,T}(\frac{t}{T}, x) \\
&= -\eta(i+1-dx)\sum_{j=1}^{d} W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})U^{d,T}(\frac{t}{T},\frac{j}{d}) \\
&\quad - \eta(dx-i)\sum_{j=1}^{d} W(\frac{t}{T},\frac{i+1}{d})W(\frac{t}{T},\frac{j}{d})U^{d,T}(\frac{t}{T},\frac{j}{d}) \\
&\quad - \eta\gamma(i+1-dx)\sum_{j=1}^{d}\big(W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d}) - \mathbb{E}\big[W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})\big]\big)\Theta(\frac{t}{T},\frac{j}{d}) \\
&\quad - \eta\gamma(dx-i)\sum_{j=1}^{d}\big(W(\frac{t}{T},\frac{i+1}{d})W(\frac{t}{T},\frac{j}{d}) - \mathbb{E}\big[W(\frac{t}{T},\frac{i+1}{d})W(\frac{t}{T},\frac{j}{d})\big]\big)\Theta(\frac{t}{T},\frac{j}{d}) \\
&\quad - \gamma\big(\Theta(\frac{t+1}{T},x) - \Theta(\frac{t}{T},x) + \eta\sum_{j=1}^{d} A(x,\frac{j}{d})\Theta(\frac{t}{T},\frac{j}{d})\big) \\
&\quad + \eta\gamma(i+1-dx)W(\frac{t}{T},\frac{i}{d})\varepsilon^t + \eta\gamma(dx-i)W(\frac{t}{T},\frac{i+1}{d})\varepsilon^t.
\end{aligned}$$

如果在等式两边都应用有界光滑测试函数 $f\colon [0,\tau] \to \mathbb{R}$，那么对任意 $s \in (0,\tau)$，有：

$$\begin{aligned}
&\sum_{t=0}^{\lfloor sT\rfloor-1} f(\frac{t}{T})\big(U^{d,T}(\frac{t+1}{T},x) - U^{d,T}(\frac{t}{T},x)\big) \\
&= -(i+1-dx)\sum_{k=1}^{3} P^s_{i,k} - (dx-i)\sum_{k=1}^{3} P^s_{i+1,k} - P^s_4,
\end{aligned} \tag{45}$$

其中，对于任意 $i \in [d]$，

$$P^s_{i,1} = \eta\sum_{t=0}^{\lfloor sT\rfloor-1}\sum_{j=1}^{d} f(\frac{t}{T})W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})U^{d,T}(\frac{t}{T},\frac{j}{d}),$$

$$P^s_{i,2} = -\eta\gamma \sum_{t=0}^{\lfloor sT\rfloor-1} f(\frac{t}{T})W(\frac{t}{T},\frac{i}{d})\varepsilon^t,$$
$$P^s_{i,3} = \eta\gamma \sum_{t=0}^{\lfloor sT\rfloor-1}\sum_{j=1}^{d} f(\frac{t}{T})\big(W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d}) - \mathbb{E}\big[W(\frac{t}{T},\frac{i}{d})W(\frac{t}{T},\frac{j}{d})\big]\big)\Theta(\frac{t}{T},\frac{j}{d}),$$
$$P^s_4 = \gamma \sum_{t=0}^{\lfloor sT\rfloor-1} f(\frac{t}{T})\big(\Theta(\frac{t+1}{T},x) - \Theta(\frac{t}{T},x) + \eta\sum_{j=1}^{d} A(x,\frac{j}{d})\Theta(\frac{t}{T},\frac{j}{d})\big).$$

与推导（34）的方式类似，我们可以将（45）的左侧写成

$$\text{LHS}(45) = -\int_0^{\frac{\lfloor sT\rfloor}{T}} U(u,x)f'(u)\mathrm{d}u + f(s)U(s,x) - f(0)U(0,x) + o(1). \tag{46}$$

接下来，我们考察（45）式的右端。右端共有 4 类项。基于第1节中引入的相同时空插值，对于每个 $l=1$、2、3、$\{P^s_{i,l}\}$，都可以插值为 $C([0,\tau];C([0,1])$ 中的函数，记为 P^d、$_l{}^T$。P^s_4 已经关于变量 x 连续，因此可以插值为 $C([0,\tau];C([0,1])$ 中的函数，记为 P^d、$_4{}^T$。根据引理 5.3，（U^d、$_T$、P^d、$_1{}^T$、P^d、$_2{}^T$、P^d、$_3{}^T$、P^d、$_4{}^T$）在 $C([0,\tau];C([0,1])^{\otimes 5}$ 中是紧的。因此，序贯弱极限在 $C([0,\tau];C([0,1])^{\otimes 5}$ 中存在。（U^d、$_T$、P^d、$_1{}^T$、P^d、$_2{}^T$、P^d、$_3{}^T$、P^d、$_4{}^T$）在 $C([0,\tau];C([0,1])^{\otimes 5}$ 中的序贯弱极限，可根据各项的弱极限推导得出。注意，对于 U^d、$_T$，其弱极限记为 U，满足一个由其他项的弱极限确定的方程，具体如下：

- 对于 P^d、$_1{}^T$，基于连续映射定理（Billingsley，2009，定理 2.7）可确定弱极限；
- 对于 P^d、$_2{}^T$，P^d、$_3{}^T$，要确定无限维过程的弱极限，只需确定其每个有限维弱极限。我们通过逐点中心极限定理确定有限维极限。
- 对于 P^d、$_4{}^T$，我们利用这样一个事实：依概率收敛到零意味着依分布收敛到零。

现在，我们来考察各项的极限。

(1) 对于与 P^s、$_1$ 相关的项，$i\in[d]$我们可以采用与处理N^s、$_1$，N^s、$_i$定理 3.3证明中类似的方法来处理。我们得到

$$-(i+1-dx)P^s_{i,1} - (dx-i)P^s_{i+1,1} = -\alpha\int_0^{\frac{\lfloor sT\rfloor}{T}}\int_0^1 f(u)A(x,y)U(u,y)\mathrm{d}y\mathrm{d}u + o(1).$$

(2) 对于与 $P^s_{i,2}$（其中某个 $i\in[d]$）相关的项，我们可以采用与定理 3.3 的证明中处理 $N^s_{i,2}$ 类似的方法。存在一个中心化高斯随机场过程 $\{\xi_2(s,x)\}_{s\in[0,\tau],x\in[0,1]}$，使得对于任意 $s_1,s_2\in[0,\tau]$ 和 x，$y\in[0,1]$，

$$\mathbb{E}[\xi_2(s_1,x)\xi_2(s_2,y)] = \sigma_2(s_1,s_2,x,y) \quad \text{with} \quad \sigma_2(s_1,s_2,x,y) = (s_1\wedge s_2)A(x,y),$$

and

$$P_{i,2}^s = \begin{cases} o(1) & \text{if } \gamma\sigma d^{-1}T^{-\frac{1}{2}} \to 0, \\ \alpha\beta \int_0^s f(u)\mathrm{d}\xi_2(u, \frac{i}{d}) + o(1) & \text{if } \gamma\sigma d^{-1}T^{-\frac{1}{2}} \to \beta \in (0, \infty). \end{cases}$$

因此

(2.1) When $\gamma\sigma d^{-1}T^{-\frac{1}{2}} \to 0$, $-(i+1-dx)P_{i,2}^s - (dx-i)P_{i+1,2}^s \to 0$ in probability.

(2.2) 当 $\gamma\sigma d^{-1}T^{-\frac{1}{2}} \to \beta \in (0, \infty)$ 时,

$$-(i+1-dx)P_{i,2}^s - (dx-i)P_{i+1,2}^s = \alpha\beta \int_0^{\frac{\lfloor sT \rfloor}{T}} f(u)\mathrm{d}\xi_2(u, x) + o(1).$$

(3) 为了研究与 $P_{i,3}$ 相关的项，其中某个 $i \in [d]$，我们先定义

$$Z_\gamma^{d,T}(\frac{t}{T}, \frac{j}{d}) := \frac{\gamma}{dT} \sum_{l=1}^{d} f(\frac{t}{T}) \big(W(\frac{t}{T}, \frac{j}{d}) W(\frac{t}{T}, \frac{l}{d}) - \mathbb{E}\big[W(\frac{t}{T}, \frac{j}{d}) W(\frac{t}{T}, \frac{l}{d}) \big] \big) \Theta(\frac{t}{T}, \frac{l}{d}),$$

对于任意 $t \le \lfloor sT \rfloor, j \in [d]$。于是我们可以写成

$$P_{i,3}^s = (\eta dT) \sum_{t=0}^{\lfloor sT \rfloor - 1} Z_\gamma^{d,T}(\frac{t}{T}, \frac{i}{d}).$$

注意，当$t_1 \neq t_2$时，$\mathbb{E}\big[Z_\gamma^{d,T}(\frac{t}{T}, \frac{i}{d})\big] = 0$和 $\{Z_\gamma^{d,T}(\frac{t_1}{T}, \frac{i}{d})\}_{i\in[d]}$与 $\{Z_\gamma^{d,T}(\frac{t_2}{T}, \frac{i}{d})\}_{i\in[d]}$相互独立。对于任意 $0 \le s_1 < s_2 \le \tau, i, j \in [d]$，我们计算得

$$\begin{aligned} &\mathbb{E}\big[P_{i,3}^{s_1} P_{j,3}^{s_2}\big] \\ =&(\eta dT)^2 \mathbb{E}\big[\sum_{t=0}^{\lfloor s_1 T \rfloor - 1} Z_\gamma^{d,T}(\frac{t}{T}, \frac{i}{d}) Z_\gamma^{d,T}(\frac{t}{T}, \frac{j}{d}) \big] \\ =&(\eta dT)^2 \frac{\gamma^2}{d^2T^2} \sum_{t=0}^{\lfloor s_1 T \rfloor - 1} \sum_{l_1, l_2 = 1}^{d} f(\frac{t}{T})^2 \Theta(\frac{t}{T}, \frac{l_1}{d}) \Theta(\frac{t}{T}, \frac{l_2}{d}) \tilde{B}(\frac{i}{d}, \frac{l_1}{d}, \frac{j}{d}, \frac{l_2}{d}) \\ =&\frac{\alpha^2\gamma^2}{T^2} \sum_{t=0}^{\lfloor s_1 T \rfloor - 1} \big(\int_0^1 \int_0^1 f(\frac{t}{T})^2 \Theta(\frac{t}{T}, x) \Theta(\frac{t}{T}, y) \tilde{B}(\frac{i}{d}, x, \frac{j}{d}, y) \mathrm{d}x\mathrm{d}y + o(1) \big), \end{aligned}$$

其中，对于任意 $x_1, x_2, x_3, x_4 \in [0,1]$，$\tilde{B}(x_1, x_2, x_3, x_4) = B(x_1, x_2, x_3, x_4) - A(x_1, x_3)A(x_2, x_4)$。第一个等式由独立性得出，最后一个等式则源于$f, \Theta, B$连续且有界这一事实。因此，我们有

$$\begin{aligned} &\mathbb{E}\big[P_{i,3}^{s_1} P_{j,3}^{s_2}\big] \\ =&\gamma^2 T^{-1} \alpha^2 \int_0^{\frac{\lfloor s_1 T \rfloor}{T}} \int_0^1 \int_0^1 f(u)^2 \Theta(u, x) \Theta(u, y) \tilde{B}(\frac{i}{d}, x, \frac{j}{d}, y) \mathrm{d}x\mathrm{d}y\mathrm{d}u + o(\gamma^2 T^{-1}). \end{aligned} \tag{47}$$

现在，我们有以下观测值。

（3.1）如果 $\gamma = o(T^{\frac{1}{2}})$，那么 $P_{i,3}^s \to 0$ 依概率且 $-(i+1-dx)P_{i,2}^s - (dx-i)P_{i+1,2}^s \to 0$ 依概率。

（3.2）如果$-\gamma T^{-\frac{1}{2}} \to \zeta \in (0,\infty)$，那么，结合（47），我们可以先对$P_{i,3}$应用林德伯格-费勒中心极限定理。由于

$$\begin{aligned}
&\mathbb{E}\big[Z_\gamma^{d,T}(\frac{t}{T},\frac{i}{d})^2\big] \\
&= (\frac{\gamma}{dT})^2 \sum_{l_1,l_2}^{d} f(\frac{t}{T})^2 \Theta(\frac{t}{T},\frac{l_1}{d})\Theta(\frac{t}{T},\frac{l_2}{d})\tilde{B}(\frac{i}{d},\frac{l_1}{d},\frac{j}{d},\frac{l_2}{d}) \\
&= \frac{\zeta^2 f(\frac{t}{T})^2}{T}\Big(\int_0^1\int_0^1 \Theta(\frac{t}{T},z_1)\Theta(\frac{t}{T},z_2)\tilde{B}(\frac{i}{d},z_1,\frac{i}{d},z_2)\mathrm{d}x\mathrm{d}y + o(1)\Big)
\end{aligned}$$

and

$$\mathrm{Var}(P_{i,3}^s) = \zeta^2\alpha^2 \int_0^{\frac{\lfloor sT\rfloor}{T}}\int_0^1\int_0^1 f(u)^2\Theta(u,z_1)\Theta(u,z_2)\tilde{B}(\frac{i}{d},x,\frac{i}{d},y)\mathrm{d}z_1\mathrm{d}z_2\mathrm{d}u + o(1).$$

当 $T\to\infty$ 时，我们有

$$\frac{\max_{0\le t\le \lfloor sT\rfloor -1}\mathbb{E}\big[Z_\gamma^{d,T}(\frac{t}{T},\frac{i}{d})^2\big]}{\mathrm{Var}(P_{i,3}^s)} \to 0.$$

因此，存在中心化高斯随机场过程 $\{\xi_3(s,x)\}_{s\in[0,\tau],x\in[0,1]}$，使得对于任意 $s_1,s_2\in[0,\tau]$ 和 x，$y\in[0,1]$，

$$\mathbb{E}[\xi_3(s_1,x)\xi_3(s_2,y)] = \sigma_3(s_1,s_2,x,y)$$

$$\text{with}\quad \sigma_3(s_1,s_2,x,y) = \int_0^{s_1\wedge s_2}\int_0^1\int_0^1 \Theta(s,z_1)\Theta(s,z_2)\tilde{B}(x,z_1,x,z_2)\mathrm{d}z_1\mathrm{d}z_2\mathrm{d}s,$$

且 $P_{i,3}^s \to \alpha\zeta\int_0^s f(u)\mathrm{d}\xi_3(u,\frac{i}{d})$ 对所有 $i\in[d]$ 依分布收敛。通过利用方程 (47) 检查随机过程 $\{(P_1^s,3,P_2^s,3,\cdots,P_d^s,3)\}_{s\ge0}$ 的协方差，我们得到 $\{(P_i^s,3)_{i\in[d]}\}_{s\ge0}\to\{(\alpha\zeta\int_0^s f(u)\mathrm{d}\xi_3(u,\frac{i}{d}))_{i\in[d]}\}_{s\ge0}$，其中极限是一个有限维高斯过程。此外，利用 $A\in C([0,1]^2)$ 和 $B(\cdot,x,\cdot,y)\in C([0,1]^2)$ 这些事实，对于任意 $z_1\in[\frac{i}{d},\frac{i+1}{d})$，$z_2\in[\frac{j}{d},\frac{j+1}{d})$，

$$\begin{aligned}
&(i+1-dz_1)(j+1-dz_2)\tilde{B}(\frac{i}{d},x,\frac{j}{d},y) \\
&+(i+1-dz_1)(dz_2-j)\tilde{B}(\frac{i}{d},x,\frac{j+1}{d},y) \\
&+(dz_1-i)(j+1-dz_2)\tilde{B}(\frac{i+1}{d},x,\frac{j+1}{d},y) \\
&+(dz_1-i)(dz_2-j)\tilde{B}(\frac{i+1}{d},x,\frac{j+1}{d},y) \\
&=\tilde{B}(z_1,x,z_2,y)+o(1)
\end{aligned}$$

并且，任何 {v1）：

$$-(i+1-dx)P_{i,3}^s - (dx-i)P_{i+1,3}^s \xrightarrow{\mathrm{d}} \alpha\zeta\int_0^s f(u)\mathrm{d}\xi_3(u,x). \tag{48}$$

还要注意，对于任意 $s\in[0,\tau]$ 和 i，$j\in[d]$，因此对于任意 $s\in[0,\tau]$ 和 x，$y\in[0,1]$，我们有 $\mathbb{E}[\xi_2(s,x)\xi_3(s,y)]=0$。因此，高斯随机场过程 $\{\xi_3(s,x)\}_{s\in[0,\tau],\ x\in[0,1]}$ 与高斯随机场过程 $\{\xi_2(s,x)\}_{s\in[0,\tau],\ x\in[0,1]}$ 相互独立。

(4) 为了研究项 P_4^s，由于 $\Theta \in C^1([0,\tau];\ C^1([0,1])$ 和 (8)，我们有

$$\begin{aligned}
&\Theta(\frac{t+1}{T},x)-\Theta(\frac{t}{T},x)\\
=&-\alpha\int_{\frac{t}{T}}^{\frac{t+1}{T}}\int_0^1 A(x,y)\Theta(s,y)\mathrm{d}y\mathrm{d}s\\
=&\frac{\alpha}{dT}\sum_{j=1}^{d}A(x,\frac{j}{d})\Theta(\frac{t}{T},\frac{j}{d})+O(d^{-1}T^{-1}+T^{-2})\\
=&\eta\sum_{j=1}^{d}\big((i+1-dx)A(\frac{i}{d},\frac{j}{d})+(dx-i)A(\frac{i+1}{d},\frac{j}{d})O(d^{-1})\big)\Theta(\frac{t}{T},\frac{j}{d})\\
&+O(d^{-1}T^{-1}+T^{-2}).
\end{aligned}$$

由于 $\eta dT\to\alpha$，我们有

$$\begin{aligned}
P_4^s&=\gamma\sum_{t=1}^{\lfloor sT\rfloor-1}f(\frac{t}{T})O(\eta+d^{-1}T^{-1}+T^{-2})\\
&=\frac{1}{T}\sum_{t=1}^{\lfloor sT\rfloor-1}f(\frac{t}{T})O(d^{-1}\gamma+T^{-1}\gamma).
\end{aligned}$$

因此，结合我们的近似结果，可得

$$\begin{aligned}
&\mathrm{RHS}(45)\\
=&-\alpha\int_0^{\frac{\lfloor sT\rfloor}{T}}\int_0^1 f(u)A(x,y)U(u,y)\mathrm{d}y\mathrm{d}u+\alpha\beta\int_0^s f(u)\mathrm{d}\xi_2(u,x)1_{\{\gamma\sigma d^{-1}T^{-\frac12}\to\beta\}}\\
&+\alpha\zeta\int_0^s f(u)\mathrm{d}\xi_3(u,x)1_{\{\gamma T^{-\frac12}\to\zeta\}}+\int_0^{\frac{\lfloor sT\rfloor}{T}}f(u)\mathrm{d}uO(\gamma d^{-1}+\gamma T^{-1})+o(1),
\end{aligned}\tag{49}$$

其中，$\{\xi_2(s,x)\}_{s\in[0,\tau],x\in[0,1]}$ 和 $\{\xi_3(s,x)\}_{s\in[0,\tau],x\in[0,1]}$ 是两个相互独立且适应于同一滤过 $\{\mathcal{F}_s\}_{s\in[0,\tau]}$ 的高斯场过程，其协方差分别由 (12) 和 (13) 给出。$\beta,\zeta\in[0,\infty)$ 因为 $\max(\gamma T^{-\frac12},\gamma d^{-1},\gamma\sigma d^{-1}T^{-\frac12})\le C_s$,2。结合 (46) 和 (49)，对于任意 $s\in[0,\tau]$，我们有

$$\begin{aligned}
&-\int_0^s U(u,x)f'(u)\mathrm{d}u+f(s)U(s,x)-f(0)U(0,x)\\
=&-\alpha\int_0^s\int_0^1 f(u)A(x,y)U(u,y)\mathrm{d}y\mathrm{d}u+\alpha\beta\int_0^s f(u)\mathrm{d}\xi_2(u,x)1_{\{\gamma\sigma d^{-1}T^{-\frac12}\to\beta\}}\\
&+\alpha\zeta\int_0^s f(u)\mathrm{d}\xi_3(u,x)1_{\{\gamma T^{-\frac12}\to\zeta\}}+O(\gamma d^{-1}+\gamma T^{-1})+o(1),
\end{aligned}$$

从而得到所需结果。 ■

定理的证明*1.10*。证明基于定理3.5中确立的结果，具体说明如下。

粒子相互作用占主导和噪声占主导的情形，分别根据定理1.10 的结论 (a) 和 (b)，通过分部积分即可得出，正如我们

在第3.2节所讨论的。定理 $C([0,\tau];\ C([0,1]))$ 的唯一性和解的存在性可由定理3.6的第 1 部分得出。

对于插值误差占主导的情形，由于 f 是光滑的，我们知道

$$\int_0^{(\cdot)} U(u,x)f'(u)\mathrm{d}u \in C^1([0,\tau]) \quad \text{and} \quad \int_0^{(\cdot)} f(u)A(x,y)U(u,y)\mathrm{d}u\mathrm{d}y \in C^1([0,\tau]).$$

因此，根据（23），f（$\cdot$）U（$\cdot$，x）$\in C^1$（$[0,\tau]$），对任意 $x\in[0,1]$，这意味着 U（$\cdot$，x）$\in C^1$（$[0,\tau]$），对任意 $x\in[0,1]$。接下来，我们可以对（23）的左侧应用分部积分。因此，U 满足（16）。

由于 A 满足假设 1.2，对任意 U_1，$U_2\in C$（$[0,\tau]$；C（$[0,1]$）），我们有

$$\begin{aligned}
&\alpha \sup_{x\in[0,1]} \left| \int_0^1 A(x,y)\left(U_1(s,y)-U_2(s,y)\right)\mathrm{d}y \right| \\
&\le \alpha \sup_{x,y\in[0,1]} |A(x,y)| \sup_{y\in[0,1]} |U_1(s,y)-U_2(s,y)| \\
&\le \alpha C_2 \sup_{y\in[0,1]} |U_1(s,y)-U_2(s,y)|.
\end{aligned}$$

因此，方程（16）右端关于 U（s，$\cdot$）在任意 $s\in[0,\tau]$下均为利普希茨连续。根据 Picard–Lindel¨of 定理（例如参见 Arnold （1992年）），对于任意初始条件 C（$[0,\tau]$；C（$[0,1]$）），方程（16）存在唯一解 U（0，$\cdot$）$=U_0$（$\cdot$）$\in C$（$[0,1]$）。

由方程 (16) 解的唯一性及定理3.5可知，$\{U^{d,T}\}_{d\ge1,T>0}$的每个子列都存在一个进一步的子列，弱收敛于唯一的 $U\in C([0,\tau];\ C([0,1]))$。因此，当 $d,T\to\infty$ 时，$\{U^{d,T}\}_{d\ge1,T>0}$弱收敛于 U。■

引理 *5.3.* 设 $P^{d,T}_{l}\in C([0,\tau];\ C([0,1]))$ 为定理 *3.5* 的证明中针对 $l=1,2,3,4$ 所定义的插值函数。在定理 *3.5* 的条件下，由$\{(U^{d,T},P^{d,T}_{1},P^{d,T}_{2},P^{d,T}_{3},P^{d,T}_{4})\}_{d\ge1,T>0}$表示的序列在空间 $C([0,\tau];\ C([0,1]))^{\otimes5}$ 中是紧的。

引理的证明 *5.3*。由于具有乘积结构，只需证明以下各项分别具有紧性：$U^{d,T}$、$P^{d,T}_{1}$、$P^{d,T}_{2}$、$P^{d,T}_{3}$、$P^{d,T}_{4}$。$\{U^{d,T}\}$的紧性已在定理 3.4 中证明。本证明中，我们证明 $\{P^{d,T}_{l}\}$在 $l=1$、2、3、4 上的紧性。

的紧性 $P^{d,T}_{1}$。首先，我们有 $P^{d,T}_{1}(0,0)=\eta f(0)\sum_{j=1}^{d}W(0,0)W(0,\frac{j}{d})U^{d,T}(0,\frac{j}{d})$。在定理 3.5 的假设下，我们有

$$\begin{aligned}
\mathbb{P}(|P_1^{d,T}(0,0)|>N) &\le N^{-2}\eta^2 f(0)^2 \sum_{j=1}^{d} B(0,\frac{j}{d},0,\frac{j}{d}) \sum_{j=1}^{d} \mathbb{E}[|U^{d,T}(0,\frac{j}{d})|^2] \\
&\le C_5 \|f\|_\infty^2 N^{-2}\eta^2 d^2 D^2.
\end{aligned}$$

由于 $\eta dT=\alpha<\infty$，我们有 $\mathbb{P}(|P^{d,T}_{1}(0,0)|>N)\lesssim N^{-2}\to0$当 $N\to\infty$时。因此，$\{P_1^{d,T}(0,0)\}_{d\ge1,T>0}$在概率空间中是紧的。接下来，我们注意到，对于任意 $0\le t_1<t_2\le\lfloor sT\rfloor-1$和 i_1，$i_2\in[d]$，

$$P_1^{d,T}(\frac{t_2}{T},\frac{i_2}{d})-P_1^{d,T}(\frac{t_1}{T},\frac{i_1}{d})$$

$$
\begin{aligned}
&= \eta \sum_{t=t_1}^{t_2-1} \sum_{j=1}^{d} f(\frac{t}{T}) W(\frac{t}{T}, \frac{i_2}{d}) W(\frac{t}{T}, \frac{j}{d}) U^{d,T}(\frac{t}{T}, \frac{j}{d}) \\
&\quad + \eta \sum_{t=0}^{t_1-1} \sum_{j=1}^{d} f(\frac{t}{T}) \big(W(\frac{t}{T}, \frac{i_2}{d}) W(\frac{t}{T}, \frac{j}{d}) - W(\frac{t}{T}, \frac{i_1}{d}) W(\frac{t}{T}, \frac{j}{d}) \big) U^{d,T}(\frac{t}{T}, \frac{j}{d}) \\
&= \eta \sum_{t=t_1}^{t_2-1} \sum_{j=1}^{d} f(\frac{t}{T}) A(\frac{i_2}{d}, \frac{j}{d}) U^{d,T}(\frac{t}{T}, \frac{j}{d}) \\
&\quad + \eta \sum_{t=0}^{t_1-1} \sum_{j=1}^{d} f(\frac{t}{T}) \big(A(\frac{i_2}{d}, \frac{j}{d}) - A(\frac{i_1}{d}, \frac{j}{d}) \big) U^{d,T}(\frac{t}{T}, \frac{j}{d}) \\
&\quad + \eta \sum_{t=t_1}^{t_2-1} \sum_{j=1}^{d} f(\frac{t}{T}) \big(W(\frac{t}{T}, \frac{i_2}{d}) W(\frac{t}{T}, \frac{j}{d}) - A(\frac{i_2}{d}, \frac{i}{d}) \big) U^{d,T}(\frac{t}{T}, \frac{j}{d}) \\
&\quad + \eta \sum_{t=0}^{t_1-1} \sum_{j=1}^{d} f(\frac{t}{T}) \Big(W(\frac{t}{T}, \frac{i_2}{d}) W(\frac{t}{T}, \frac{j}{d}) - W(\frac{t}{T}, \frac{i_1}{d}) W(\frac{t}{T}, \frac{j}{d}) \\
&\qquad\qquad - A(\frac{i_2}{d}, \frac{j}{d}) + A(\frac{i_1}{d}, \frac{j}{d})) \Big) U^{d,T}(\frac{t}{T}, \frac{j}{d}).
\end{aligned}
$$

由命题 $\mathbb{E}[M_{i,1}^4]$, $\mathbb{E}[M_{i,2}^4]$和 $\mathbb{E}[M_{i,j,1}^4]$, $\mathbb{E}[M_{i,j,2}^4]$的证明中对 5.1 和命题5.2 的界估计论证可知,

$$
\begin{aligned}
&\mathbb{E}[|P_1^{d,T}(\frac{t_2}{T}, \frac{i_2}{d}) - P_1^{d,T}(\frac{t_1}{T}, \frac{i_1}{d})|^4] \\
\lesssim\ & \eta^4 \|f\|_\infty^4 d^4 (t_2 - t_1)^4 + \eta^4 \|f\|_\infty^4 d^4 (t_2 - t_1)^2 \\
&\quad + \eta^4 \|f\|_\infty^4 (i_2 - i_1)^4 t_1^4 + \eta^4 \|f\|_\infty^4 (i_2 - i_1)^4 t_1^2 \\
\lesssim\ & (\eta d T)^4 \big(\frac{t_2 - t_1}{T}\big)^2 + (\eta d T)^4 \big(\frac{i_2 - i_1}{d}\big)^4.
\end{aligned}
$$

由于 $P^{d,\cdot,1}{}^T$关于两个变量均为分段线性的，且对所有 $d \geq 1$, $T > 0$、所有 $0 \leq s_1 < s_2 \leq \tau$ 以及 $0 \leq x_1 < x_2 \leq 1$, 有

$$
\mathbb{E}\big[\big|P_1^{d,T}(s_1, x_1) - P_1^{d,T}(s_2, x_2)\big|^4\big] \lesssim (s_2 - s_1)^2 + (x_1 - x_2)^4.
$$

$P^{d,\cdot,1}{}^T$在 $C([0,\tau]; C([0,1]))$ 中的紧性由柯尔莫哥洛夫紧性判据推出。

的紧性 $P^{d,\cdot,2}{}^T$。首先，我们有 $P^{d,\cdot,2}{}^T(0,\ 0) = -\eta\gamma f(0) W(0,\ 0)\varepsilon^0$。在定理3.5的假设下，我们有

$$
\begin{aligned}
\mathbb{P}(|P_2^{d,T}(0,0)| > N) &\leq N^{-2} \eta^2 \gamma^2 \sigma^2 f(0)^2 A(0,0) \\
&\leq C_2 \|f\|_\infty^2 N^{-2} \eta^2 \gamma^2 \sigma^2.
\end{aligned}
$$

由于 $\eta d T = \alpha$和 $\gamma \sigma d^{-1} T^{-\frac{1}{2}} \leq C_{s,2}$, $\mathbb{P}(|P^{d,\cdot,2}{}^T(0,0)| > N) \lesssim N^{-2} \to 0$当 $N \to \infty$时。因此，$\{P_2^{d,T}(0,0)\}_{d \geq 1, T > 0}$在该概率空间中是紧的。接下来，我们注意到，对于任意 $0 \leq t_1 < t_2 \leq \lfloor sT \rfloor - 1$和 i_1, $i_2 \in [d]$,

$$
P_2^{d,T}(\frac{t_2}{T}, \frac{i_2}{d}) - P_2^{d,T}(\frac{t_1}{T}, \frac{i_1}{d})
$$

$$= -\eta\gamma \sum_{t=t_1}^{t_2-1} f(\frac{t}{T}) W(\frac{t}{T}, \frac{i_2}{d}) \varepsilon^t - \eta\gamma \sum_{t=0}^{t_1-1} \sum_{j=1}^{d} f(\frac{t}{T}) (W(\frac{t}{T}, \frac{i_2}{d}) - W(\frac{t}{T}, \frac{i_1}{d})) \varepsilon^t.$$

由命题 $\mathbb{E}[M_{i,5}^4]$ 和 $\mathbb{E}[M_{i,j,5}^4]$,5.1 和命题 5.2 证明中对 进行界定的论证可知

$$\begin{aligned}
&\mathbb{E}[|P_2^{d,T}(\frac{t_2}{T}, \frac{i_2}{d}) - P_2^{d,T}(\frac{t_1}{T}, \frac{i_1}{d})|^4] \\
&\lesssim \gamma^4\eta^4 \|f\|_\infty^4 \sigma^4 (t_2 - t_1)^2 + \gamma^4\eta^4 \|f\|_\infty^4 \sigma^4 T^2 \big(\frac{i_2 - i_1}{d}\big)^4 \big(\frac{t_1}{T}\big)^2 \\
&\lesssim (\eta\gamma\sigma T^{\frac{1}{2}})^4 \big(\frac{t_2 - t_1}{T}\big)^2 + (\eta\gamma\sigma T^{\frac{1}{2}})^4 \big(\frac{i_2 - i_1}{d}\big)^4.
\end{aligned}$$

由于 $P^{d,\ 2}{}^T$关于两个变量都是分段线性的，并且对所有 $d \geq 1$，$T > 0$以及所有 $0 \leq s_1 < s_2 \leq \tau$ 和 $0 \leq x_1 < x_2 \leq 1$，$\eta\gamma\sigma T^{-\frac{1}{2}} = \alpha(\eta\gamma d^{-1} T^{-\frac{1}{2}}) \leq \alpha C_{s,\ 2}$，我们有

$$\mathbb{E}\big[\big|P_1^{d,T}(s_1, x_1) - P_1^{d,T}(s_2, x_2)\big|^4\big] \lesssim (s_2 - s_1)^2 + (x_1 - x_2)^4.$$

紧　性 $P^{d,\ 2}{}^T$在 $C([0,\tau]; C([0,1])$中的紧性由柯尔莫哥洛夫紧性判据可得　　.

的紧性 $P^{d,\ 3}{}^T$。注意，在$(t,\ x)$处，$=(0,\ 0)$，我们有

$$P_3^{d,T}(0,0) = \eta\gamma f(0) \sum_{j=1}^{d} \big(W(0,0) W(0, \frac{j}{d}) - A(0, \frac{j}{d})\big) \Theta(0, \frac{j}{d}).$$

在定理3.5的假设下，我们有

$$\begin{aligned}
\mathbb{P}(|P_3^{d,T}(0,0)| > N) &\leq N^{-2}\eta^2\gamma^2 f(0)^2 \big(\sum_{j=1}^{d} B(0, \frac{j}{d}, 0, \frac{j}{d}) - A(0, \frac{j}{d})^2\big) \big(\sum_{j=1}^{d} \mathbb{E}[|\Theta(0, \frac{j}{d})|^2]\big) \\
&\leq (C_5 + C_2^2) \|f\|_\infty^2 \|\Theta(0,\cdot)\|_\infty^2 N^{-2}\eta^2\gamma^2 d^2.
\end{aligned}$$

由于 $\eta dT = \alpha$和 $\gamma T^{-\frac{1}{2}} \leq C_{s,\ 2}$，$\mathbb{P}(|P^{d,\ 3}{}^T(0,\ 0)| > N) \lesssim N^{-2} \to 0$当 $N \to \infty$时。因此，$\{P_3^{d,T}(0,\ 0)\}_{d \geq 1,\ T>0}$在该概率空间中是紧的。接下来，我们注意到，对于任意 $0 \leq t_1 < t_2 \leq \lfloor sT \rfloor - 1$和 i_1，$i_2 \in [d]$,

$$\begin{aligned}
&P_3^{d,T}(\frac{t_2}{T}, \frac{i_2}{d}) - P_3^{d,T}(\frac{t_1}{T}, \frac{i_1}{d}) \\
&= \gamma\eta \sum_{t=t_1}^{t_2-1} \sum_{j=1}^{d} f(\frac{t}{T}) \big(W(\frac{t}{T}, \frac{i_2}{d}) W(\frac{t}{T}, \frac{j}{d}) - A(\frac{i_2}{d}, \frac{i}{d})\big) \Theta(\frac{t}{T}, \frac{j}{d}) \\
&+ \gamma\eta \sum_{t=0}^{t_1-1} \sum_{j=1}^{d} f(\frac{t}{T}) \Big(W(\frac{t}{T}, \frac{i_2}{d}) W(\frac{t}{T}, \frac{j}{d}) - W(\frac{t}{T}, \frac{i_1}{d}) W(\frac{t}{T}, \frac{j}{d}) \\
&\qquad - A(\frac{i_2}{d}, \frac{j}{d}) + A(\frac{i_1}{d}, \frac{j}{d})\Big) \Theta(\frac{t}{T}, \frac{j}{d}).
\end{aligned}$$

由命题 $\mathbb{E}[M_{i,3}^4]$ 和 $\mathbb{E}[M_{i,j,4}^4]$ 的界定论证可知，在命题 5.1和命题 5.2的证明中,

$$\mathbb{E}[|P_3^{d,T}(\frac{t_2}{T}, \frac{i_2}{d}) - P_3^{d,T}(\frac{t_1}{T}, \frac{i_1}{d})|^4]$$

$$\lesssim \gamma^4\eta^4 \|f\|_\infty^4 d^4(t_2-t_1)^2 + \gamma^4\eta^4 \|f\|_\infty^4 d^4T^2\big(\frac{i_2-i_1}{d}\big)^4\big(\frac{t_1}{T}\big)^2$$
$$\lesssim (\eta\gamma dT^{\frac{1}{2}})^4\big(\frac{t_2-t_1}{T}\big)^2 + (\eta\gamma dT^{\frac{1}{2}})^4\big(\frac{i_2-i_1}{d}\big)^4.$$

由于 $P^{d,\ 3}{}^T$关于两个变量均为分段线性函数，且 $\eta\gamma dT^{-\frac{1}{2}} = \alpha(\gamma T^{-\frac{1}{2}}) \le \alpha C_{s,\ 2}$对所有 $d \ge 1$, $T > 0$以及所有 $0 \le s_1 < s_2 \le \tau$和 $0 \le x_1 < x_2 \le 1$均成立，因此我们有

$$\mathbb{E}\big[\big|P_3^{d,T}(s_1,x_1) - P_3^{d,T}(s_2,x_2)\big|^4\big] \lesssim (s_2-s_1)^2 + (x_1-x_2)^4.$$

紧性 $P^{d,\ 3}{}^T$在 $C([0,\tau];\ C([0,1]))$ 中的紧性可由柯尔莫哥洛夫紧性判据推出 .

的紧性$_4{}^T$。首先，我们有 $P^{d,\ 4}{}^T(0,\ 0) = \gamma\eta f(0)\sum_{j=1}^d A(0,\ \frac{j}{d})\Theta(0,\ \frac{j}{d})$。根据定理3.5中的假设，我们有

$$\mathbb{P}(|P_4^{d,T}(0,0)| > N) \le N^{-2}\eta^2\gamma^2 f(0)^2\big(\sum_{j=1}^d A(0,\frac{j}{d})^2\big)\big(\sum_{j=1}^d \mathbb{E}[|\Theta(0,\frac{j}{d})|^2]\big)$$
$$\le C_2^2 \|f\|_\infty^2 \|\Theta(0,\cdot)\|_\infty^2 N^{-2}\eta^2\gamma^2 d^2.$$

由于 $\eta dT = \alpha$和 $\gamma T^{-\frac{1}{2}} \lesssim C_{s,\ 2}$, $\mathbb{P}(|P^{d,\ 4}{}^T(0,\ 0)| > N) \lesssim N^{-2} \to 0$当 $N \to \infty$时。因此，$\{P_4^{d,T}(0,\ 0)\}_{d\ge 1,\ T>0}$在该概率空间中是紧的。接下来，我们注意到，对于任意 $0 \le t_1 < t_2 \le \lfloor sT\rfloor - 1$和 $i_1,\ i_2 \in [d]$,

$$\begin{aligned}
&P_4^{d,T}(\frac{t_2}{T},\frac{i_2}{d}) - P_4^{d,T}(\frac{t_1}{T},\frac{i_1}{d})\\
&= \gamma\sum_{t=t_1}^{t_2-1} f(\frac{t}{T})\big(\Theta(\frac{t+1}{T},\frac{i_2}{d}) - \Theta(\frac{t}{T},\frac{i_2}{d}) + \eta\sum_{j=1}^d A(\frac{i_2}{d},\frac{j}{d})\Theta(\frac{t}{T},\frac{j}{d})\big)\\
&\quad + \gamma\sum_{t=0}^{t_1-1} f(\frac{t}{T})\Big(\Theta(\frac{t+1}{T},\frac{i_2}{d}) - \Theta(\frac{t}{T},\frac{i_2}{d}) - \Theta(\frac{t+1}{T},\frac{i_1}{d}) + \Theta(\frac{t}{T},\frac{i_1}{d})\\
&\qquad\qquad + \eta\sum_{j=1}^d \big(A(\frac{i_2}{d},\frac{j}{d}) - A(\frac{i_1}{d},\frac{j}{d})\big)\Theta(\frac{t}{T},\frac{j}{d})\Big).
\end{aligned}$$

由命题 $\mathbb{E}[M_{i,4}^4]$ 和 $\mathbb{E}[M_{i,j,3}^4]$的证明中对5.1和命题5.2的界定论证可知

$$\begin{aligned}
&\mathbb{E}[|P_4^{d,T}(\frac{t_2}{T},\frac{i_2}{d}) - P_4^{d,T}(\frac{t_1}{T},\frac{i_1}{d})|^4]\\
&\lesssim \gamma^4\eta^4\|f\|_\infty^4(t_2-t_1)^4 + \gamma^4\eta^4\|f\|_\infty^4 d^4T^{-4}(t_2-t_1)^4\\
&\quad + \gamma^4T^{-4}\|f\|_\infty^4\big(\frac{i_2-i_1}{d}\big)^4\big(\frac{t_1}{T}\big)^4 + \gamma^4 d^{-4}\|f\|_\infty^4\big(\frac{i_2-i_1}{d}\big)^4\big(\frac{t_1}{T}\big)^4\\
&\lesssim (\gamma d^{-1})^4\big(\frac{t_2-t_1}{T}\big)^4 + (\gamma T^{-1})^4\big(\frac{i_2-i_1}{d}\big)^4.
\end{aligned}$$

由于 $P^{d,\ 4}{}^T$关于两个变量均为分段线性函数，且对所有 $d \ge 1$, $T > 0$、所有$0 \le s_1 < s_2 \le \tau$和 $0 \le x_1 < x_2 \le 1$，都有 γd^{-1}, $\gamma T^{-\frac{1}{2}} \le C_{s,\ 2}$

$$\mathbb{E}\big[\big|P_4^{d,T}(s_1,x_1) - P_4^{d,T}(s_2,x_2)\big|^4\big] \lesssim (s_2-s_1)^4 + (x_1-x_2)^4.$$

紧性 在 $C([0,\tau];\ C([0,1]))$ 中 $P^{d,\ 4}{}^T$ 的紧性可由柯尔莫哥洛夫紧性判据推出 . ■

6 随机微分方程存在性与唯一性的证明

我们关于随机微分方程存在性和唯一性的证明依赖于以下性质： ξ_1、 ξ_2、 ξ_3 几乎必然有界且连续，这一点可通过应用（Adler andTaylor， 2007， 定理 1.4.1）来验证。在以下证明中，为简洁起见，我们将使用这一性质，而不再反复提及。

定理 *3.6 (a)* 部分的证明。 **存在性：** 我们基于皮卡迭代论证证明存在性。令 $U_0(s,x) = U(0,x)$ 对所有 $s \in [0,\tau]$成立。皮卡迭代定义如下：

$$\begin{aligned} U_k(s,x) = U(0,x) - \alpha \int_0^s \int_0^1 A(x,y) U_{k-1}(u,y) \mathrm{d}y \mathrm{d}u \\ + \alpha\beta \int_0^s \mathrm{d}\xi_2(u,x) + \alpha\zeta \int_0^s \mathrm{d}\xi_3(u,x). \end{aligned} \tag{50}$$

根据 ξ_2， ξ_3 的定义，若 $U(0,\cdot) \in C([0,1])$，则对所有 $s \in [0,\tau]$ 和 $k \geq 0$， $U_k(s,\cdot) \in C([0,1])$。接下来证明 $U_k \in C([0,\tau]; C([0,1]))$。对任意 $K > 0$，定义停时：

$$\tau_{K,\Upsilon} := \min(\tau, \inf\{s \in [0,\tau] : \max_{0\leq k\leq K} \|U_k(s,\cdot)\|_\infty \geq \Upsilon\}).$$

容易看出，当 $\Upsilon \to \infty$ 时， $\tau_{K,\Upsilon} \to \tau$ 几乎处处成立。定义 $U_k^\Upsilon(s,x) := U_k(s \wedge \tau_{K,\Upsilon}, x)$。我们有

$$\begin{aligned} U_k^\Upsilon(s,x) = U(0,x) - \alpha \int_0^s \int_0^1 A(x,y) U_{k-1}(u \wedge \tau_\Upsilon, y) 1_{(0,\tau_\Upsilon)}(u) \mathrm{d}y \mathrm{d}u \\ + \alpha\beta \int_0^s 1_{(0,\tau_\Upsilon)}(u) \mathrm{d}\xi_2(u \wedge \tau_\Upsilon, x) + \alpha\zeta \int_0^s 1_{(0,\tau_\Upsilon)}(u) \mathrm{d}\xi_3(u \wedge \tau_\Upsilon, x). \end{aligned}$$

因此，在假设1.1下，我们得到

$$\begin{aligned} &\max_{0\leq k\leq K} \|U_k^\Upsilon(s,\cdot)\|_\infty \\ &\leq \|U(0,\cdot)\|_\infty + C_2\alpha \int_0^s \int_0^1 \max_{0\leq k\leq K} \|U_k^\Upsilon(u,\cdot)\|_\infty \mathrm{d}y \mathrm{d}u \\ &\quad + \alpha\beta \sup_{x\in[0,1]} \Big| \int_0^s 1_{(0,\tau_\Upsilon)}(u) \mathrm{d}\xi_2(u,x) \Big| + \alpha\zeta \sup_{x\in[0,1]} \Big| \int_0^s 1_{(0,\tau_\Upsilon)}(u) \mathrm{d}\xi_3(u,x) \Big|. \end{aligned}$$

取期望，我们得到

$$\begin{aligned} &\mathbb{E}\big[\max_{0\leq k\leq K} \sup_{s\in[0,\tau]} \|U_k^\Upsilon(s,\cdot)\|_\infty \big] \\ &\leq \mathbb{E}\big[\|U(0,\cdot)\|_\infty \big] + C_2\alpha \int_0^\tau \mathbb{E}\big[\max_{0\leq k\leq K} \|U_k^\Upsilon(u,\cdot)\|_\infty \big] \mathrm{d}u \\ &\quad + \alpha\beta \mathbb{E}\big[\sup_{x\in[0,1]} \Big(\int_0^\tau 1_{(0,\tau_{K,\Upsilon})}(u) \mathrm{d}\xi_2(u,x) \Big)^2 \big]^{\frac{1}{2}} \\ &\quad + \alpha\zeta \mathbb{E}\big[\sup_{x\in[0,1]} \Big(\int_0^\tau 1_{(0,\tau_{K,\Upsilon})}(u) \mathrm{d}\xi_3(u,x) \Big)^2 \big]^{\frac{1}{2}}. \end{aligned}$$

根据（12）和（13），在假设1.1下，我们有

$$\mathbb{E}\big[\max_{0\leq k\leq K} \sup_{s\in[0,\tau]} \|U_k^\Upsilon(s,\cdot)\|_\infty \big]$$

$$
\begin{aligned}
\leq & \mathbb{E}\big[\,\|U(0,\cdot)\|_\infty\,\big] + C_2\alpha\int_0^\tau \mathbb{E}\big[\max_{0\leq k\leq K}\big\|U_k^\Upsilon(u,\cdot)\big\|_\infty\big]\mathrm{d}u \\
& + \alpha\beta\mathbb{E}\big[\sup_{x\in[0,1]}\int_0^\tau A(x,x)\mathrm{d}(u\wedge\tau_{K,\Upsilon})\big]^{\frac{1}{2}} \\
& + \alpha\zeta\mathbb{E}\big[\sup_{x\in[0,1]}\int_0^\tau\int_0^1\int_0^1 \Theta(u,z_1)\Theta(u,z_2)\tilde{B}(x,z_1,x,z_2)\mathrm{d}z_1\mathrm{d}z_2\mathrm{d}(u\wedge\tau_{K,\Upsilon})\big]^{\frac{1}{2}} \\
\leq & \mathbb{E}\left[\|U(0,\cdot)\|_\infty\right] + C_2\alpha\int_0^\tau \mathbb{E}\big[\max_{0\leq k\leq K}\sup_{u\in[0,s]}\big\|U_k^\Upsilon(u,\cdot)\big\|_\infty\big]\mathrm{d}s \\
& + C_2^{\frac{1}{2}}\alpha\beta\tau^{\frac{1}{2}} + (C_2^2+C_5)^{\frac{1}{2}}\alpha\zeta\left\|\Theta\right\|_\infty\tau^{\frac{1}{2}},
\end{aligned}
$$

其中 $\tilde{B}(x,\ z_1,\ x,\ z_2) = B(x,\ z_1,\ x,\ z_2) - A(x,\ z_1)A(x,\ z_2)$。利用格朗沃尔不等式，我们得到

$$
\begin{aligned}
& \mathbb{E}\big[\max_{0\leq k\leq K}\sup_{s\in[0,\tau]}\big\|U_k^\Upsilon(s,\cdot)\big\|_\infty\big] \\
\leq & \left(\mathbb{E}\big[\,\|U(0,\cdot)\|_\infty\,\big] + C_2^{\frac{1}{2}}\alpha\beta\tau^{\frac{1}{2}} + (C_5+C_2^2)^{\frac{1}{2}}\alpha\zeta\left\|\Theta\right\|_\infty\tau^{\frac{1}{2}}\right)\exp(C_2\alpha\tau).
\end{aligned}
$$

由于 K 是任意的，我们可以令 $K,\ \Upsilon\to\infty$，根据单调收敛定理，对于任意 k，我们证明

$$
\begin{aligned}
& \mathbb{E}\big[\sup_{s\in[0,\tau]}\|U_k(s,\cdot)\|_\infty\big] \\
\leq & \left(\mathbb{E}\big[\,\|U(0,\cdot)\|_\infty\,\big] + C_2^{\frac{1}{2}}\alpha\beta\tau^{\frac{1}{2}} + (C_5+C_2^2)^{\frac{1}{2}}\alpha\zeta\left\|\Theta\right\|_\infty\tau^{\frac{1}{2}}\right)\exp(C_2\alpha\tau).
\end{aligned}
$$

因此，对于任意 $k,\ U_k\in C\left([0,\tau];\ C([0,1])\right)$。接下来，我们证明 $\{U_k\}_{k=1}^\infty$ 在空间中收敛$C([0,\tau];\ C([0,1]))$。由式（50），对于任意 $k\geq 1$和 $s\in[0,\tau]$,

$$
\begin{aligned}
\|U_{k+1}(s,\cdot)-U_k(s,\cdot)\|_\infty &= \alpha\left\|\int_0^s\int_0^1 A(\cdot,y)\left(U_k(u,y)-U_{k-1}(u,y)\right)\mathrm{d}y\mathrm{d}u\right\|_\infty \\
&\leq C_2\alpha s\int_0^s\|U_k(u,\cdot)-U_{k-1}(u,\cdot)\|_\infty\,\mathrm{d}u,
\end{aligned}
$$

and

$$
\begin{aligned}
\|U_1(s,\cdot)-U_0(s,\cdot)\|_\infty \leq & \alpha\left\|\int_0^s\int_0^1 A(\cdot,y)U_0(y)\mathrm{d}y\mathrm{d}u\right\|_\infty \\
& + \alpha\beta\sup_{x\in[0,1]}\Big|\int_0^s\mathrm{d}\xi_2(u,x)\Big| + \alpha\zeta\sup_{x\in[0,1]}\Big|\int_0^s\mathrm{d}\xi_3(u,x)\Big|.
\end{aligned}
$$

因此，对于任意 $s\in[0,\tau]$,

$$
\begin{aligned}
& \mathbb{E}\big[\sup_{r\in[0,s]}\|U_1(r,\cdot)-U_0(r,\cdot)\|_\infty\big] \\
\leq & (1+C_2\alpha s)\mathbb{E}\big[\,\|U(0,\cdot)\|_\infty\,\big] + C_2^{\frac{1}{2}}\alpha\beta s^{\frac{1}{2}} + (C_5+C_2^2)^{\frac{1}{2}}\alpha\zeta\left\|\Theta\right\|_\infty s^{\frac{1}{2}} \\
:= & C(s)\leq C(\tau)<\infty,
\end{aligned}
$$

and

$$
\mathbb{E}\big[\sup_{r\in[0,s]}\|U_{k+1}(r,\cdot)-U_k(r,\cdot)\|_\infty\big]
$$

$$\leq C_2\alpha \int_0^s \mathbb{E}\big[\sup_{u\in[0,r]} \|U_k(u,\cdot) - U_{k-1}(u,\cdot)\|_\infty\big]\mathrm{d}u.$$

通过归纳法，我们得到对于任意 $k \geq 1$,

$$\mathbb{E}\big[\sup_{r\in[0,s]} \|U_{k+1}(r,\cdot) - U_k(r,\cdot)\|_\infty\big] \leq \frac{C(\tau)\left(C_2\alpha s\right)^k}{k!}.$$

因此，根据马尔可夫不等式,

$$\mathbb{P}\big(\sup_{r\in[0,s]} \|U_{k+1}(r,\cdot) - U_k(r,\cdot)\|_\infty > 2^{-k}\big) \leq \frac{C(\tau)\left(2C_2\alpha s\right)^k}{k!}.$$

设 Ω 为高斯场过程 ξ_2 and ξ_3 所定义的路径空间。根据 Borel-Cantelli 引理，对于几乎所有 $\omega \in \Omega$，存在 $k(\omega)$，使得对任意 $k \geq k(\omega)$，均有 $\sup_{r\in[0,s]} \|U_{k+1}(r,\cdot) - U_k(r,\cdot)\|_\infty \leq \frac{1}{2^k}$。因此，以概率 1，$\{U_k(\cdot,\cdot)\}_{k\geq 1}$ 在 $C([0,\tau]; \overline{C}([0,1]))$ 中收敛，其极限为 $U(\cdot,\cdot) \in C([0,\tau]; C([0,1]))$。此外，通过对 (50) 取极限，可以验证 U 满足随机微分方程 (14)。因此，存在性得证。

唯一性：假设存在两个解， U， U 满足随机微分方程（14）¯，由（51）可得，对任意 $s \in [0,\tau]$, $x \in [0,1]$,

$$\overline{U}(s,x) - U(s,x) = -\alpha \int_0^s \int_0^1 A(x,y)\left(\overline{U}(u,y) - U(u,y)\right)\mathrm{d}y\mathrm{d}u,$$

这意味着

$$\mathbb{E}\big[\sup_{s\in[0,\tau]} \left\|\overline{U}(s,\cdot) - U(s,\cdot)\right\|_\infty\big] \leq C_2\alpha\mathbb{E}\big[\int_0^\tau \sup_{r\in[0,s]} \left\|\overline{U}(r,\cdot) - U(r,\cdot)\right\|_\infty \mathrm{d}s\big].$$

由格朗沃尔不等式可得

$$\mathbb{E}\big[\sup_{s\in[0,\tau]} \left\|\overline{U}(s,\cdot) - U(s,\cdot)\right\|_\infty\big] \leq \mathbb{E}\big[\left\|\overline{U}(0,\cdot) - U(0,\cdot)\right\|_\infty\big]\exp\left(C_2\alpha\tau\right) = 0.$$

因此，在 14 中，存在唯一的解 $C([0,\tau]; C([0,1]))$。 ■

定理 *3.6* 的（*b*）部分的证明。 **稳定性：** 假设随机微分方程（14）存在解，则积分形式解可写为 $\forall\, s \in [0,\tau]$

$$\begin{aligned} U(s,x) = U(0,x) &- \alpha \int_0^s \int_0^1 A(x,y)U(u,y)\mathrm{d}y\mathrm{d}u \\ &+ \alpha\beta \int_0^s \mathrm{d}\xi_2(u,x) + \alpha\zeta \int_0^s \mathrm{d}\xi_3(u,x). \end{aligned} \tag{51}$$

定义停时 $\tau_N := \min(\tau,\ \inf\{s \in [0,\tau]:\ \|U(s,\ \cdot)\|_{L^2([0,1])} \geq N\})$。容易看出，$\tau_N \to \tau$几乎处处，当 $N \to \infty$ 时。定义 $U^N(s,\ x) := U(s \wedge \tau_N,\ x)$。我们有

$$\begin{aligned} U^N(s,x) = U(0,x) &- \alpha \int_0^s \int_0^1 A(x,y)U(u \wedge \tau_N, y)1_{(0,\tau_N)}(u)\mathrm{d}y\mathrm{d}u \\ &+ \alpha\beta \int_0^s 1_{(0,\tau_N)}(u)\mathrm{d}\xi_2(u \wedge \tau_N, x) + \alpha\zeta \int_0^s 1_{(0,\tau_N)}(u)\mathrm{d}\xi_3(u \wedge \tau_N, x). \end{aligned}$$

因此，我们在 $L^2([0,1])$ 空间中得到了对 $U^N(s,\ \cdot)$ 的估计，

$$\int_0^1 \left|U^N(s,x)\right|^2 \mathrm{d}x \le 4\left\|U(0,\cdot)\right\|^2_{L^2([0,1])} + 4\alpha^2 s \int_0^s \int_0^1 \Big(\int_0^1 A(x,y)U^N(u,y)\mathrm{d}y\Big)^2 \mathrm{d}x\mathrm{d}u$$
$$+ 4\alpha^2\beta^2 \int_0^1 \Big|\int_0^s 1_{(0,\tau_N)}(u)\mathrm{d}\xi_2(u,x)\Big|^2 \mathrm{d}x + 4\alpha^2\zeta^2 \int_0^1 \Big|\int_0^s 1_{(0,\tau_N)}(u)\mathrm{d}\xi_3(u,x)\Big|^2 \mathrm{d}x.$$

在假设 1.2下，对于积分算子，我们有

$$\mathcal{A} : g \in L^2([0,1]) \mapsto \int_0^1 A(x,y)g(y)\mathrm{d}y \in L^2([0,1]),$$

$$\|\mathcal{A}\|_{op} \coloneqq \sup_{\|g\|_{L^2([0,1])}=1} \int_0^1 \Big(\int_0^1 A(x,y)g(y)\mathrm{d}y\Big)^2 \mathrm{d}x \le C_2.$$

因此可得

$$\left\|U^N(s,\cdot)\right\|^2_{L^2([0,1])} \le 4\left\|U(0,\cdot)\right\|^2_{L^2([0,1])} + 4C_2^2\alpha^2 s \int_0^s \left\|U^N(u,\cdot)\right\|^2_{L^2([0,1])} \mathrm{d}s$$
$$+ 4\alpha^2\beta^2 \int_0^1 \left(\int_0^s 1_{(0,\tau_N)}(u)\mathrm{d}\xi_2(u,x)\right)^2 \mathrm{d}x + 4\alpha^2\zeta^2 \int_0^1 \left(\int_0^s 1_{(0,\tau_N)}(u)\mathrm{d}\xi_3(u,x)\right)^2 \mathrm{d}x.$$

对 $[0,\tau]$ 取上确界并取期望，可得

$$\mathbb{E}\big[\sup_{s\in[0,\tau]} \left\|U^N(s,\cdot)\right\|^2_{L^2([0,1])}\big]$$
$$\le 4\mathbb{E}\big[\left\|U(0,\cdot)\right\|^2_{L^2([0,1])}\big] + 4C_2^2\alpha^2\tau \int_0^\tau \mathbb{E}\big[\left\|U^N(s,\cdot)\right\|^2_{L^2([0,1])}\big]\mathrm{d}s$$
$$+ 4\alpha^2\beta^2 \sup_{s\in[0,\tau]} \int_0^1 \mathbb{E}\Big[\Big(\int_0^s 1_{(0,\tau_N)}(u)\mathrm{d}\xi_2(u,x)\Big)^2\Big]\mathrm{d}x$$
$$4\alpha^2\zeta^2 \sup_{s\in[0,\tau]} \int_0^1 \mathbb{E}\Big[\Big(\int_0^s 1_{(0,\tau_N)}(u)\mathrm{d}\xi_3(u,x)\Big)^2\Big]\mathrm{d}x.$$

注意，对于任意固定的 $x \in [0,1]$，$\{\xi_i(s,\ x)\}_{s\in[0,\tau]}$ 是定义在 $[0,\tau]$ 上的高斯过程，其中 $i = 2, 3$. 根据（12）和（13），$\{\xi_2(s,\ x)\}_{s\in[0,\tau]}$ 和 $\{\xi_3(s,\ x)\}_{s\in[0,\tau]}$ 的二次变差分别为

$$[\mathrm{d}\xi_2(\cdot,x)]_s = A(x,x)s,$$
$$[\mathrm{d}\xi_3(\cdot,x)]_s = \int_0^1 \int_0^1 \Theta(s,z_1)\Theta(s,z_2)\tilde{B}(x,z_1,x,z_2)\mathrm{d}z_1\mathrm{d}z_2,$$

其中 $\tilde{B}(x,z_1,x,z_2) = B(x,x_1,x,z_2) - A(x,z_1)A(x,z_2)$。根据假设 1.2 以及伊藤公式（Revuzand Yor，1999年，定理 3.3），可得

$$\mathbb{E}\big[\sup_{s\in[0,\tau]} \left\|U^N(s,\cdot)\right\|^2_{L^2([0,1])}\big]$$
$$\le 4\mathbb{E}\big[\left\|U(0,\cdot)\right\|^2_{L^2([0,1])}\big] + 4C_2^2\alpha^2\tau \int_0^\tau \mathbb{E}\big[\left\|U^N(s,\cdot)\right\|^2_{L^2([0,1])}\big]\mathrm{d}s$$

$$
\begin{aligned}
&+4\alpha^2\beta^2\int_0^1\int_0^\tau A(x,x)\mathrm{d}(s\wedge\tau_N)\mathrm{d}x\\
&+4\alpha^2\zeta^2\int_0^1\int_0^\tau\int_0^1\int_0^1\Theta(s,z_1)\Theta(s,z_2)\tilde{B}(x,z_1,x,z_2)\mathrm{d}z_1\mathrm{d}z_2 d(s\wedge\tau_N)\mathrm{d}x\\
\le\ & 4\mathbb{E}\big[\,\|U(0,\cdot)\|^2_{L^2([0,1])}\,\big]+4C_2^2\alpha^2\tau\int_0^\tau\mathbb{E}\big[\,\big\|U^N(s,\cdot)\big\|^2_{L^2([0,1])}\,\big]\mathrm{d}s\\
&+4C_2\alpha^2\beta^2\tau+4\left(C_5+C_2^2\right)\alpha^2\zeta^2\int_0^\tau\|\Theta(s,\cdot)\|^2_{L^2([0,1])}\,\mathrm{d}s\\
\le\ & 4\mathbb{E}\big[\,\|U(0,\cdot)\|^2_{L^2([0,1])}\,\big]+4C_2^2\alpha^2\tau\int_0^\tau\mathbb{E}\big[\sup_{r\in[0,s]}\big\|U^N(r,\cdot)\big\|^2_{L^2([0,1])}\,\big]\mathrm{d}s\\
&+4C_2\alpha^2\beta^2\tau+4(C_5+C_2^2)\alpha^2\zeta^2\int_0^\tau\|\Theta(s,\cdot)\|^2_{L^2([0,1])}\,\mathrm{d}s.
\end{aligned}
$$

由格朗沃尔不等式可得

$$
\begin{aligned}
&\mathbb{E}\big[\sup_{s\in[0,\tau]}\big\|U^N(s,\cdot)\big\|^2_{L^2([0,1])}\,\big]\\
\le& 4\Bigg(\mathbb{E}\big[\,\|U(0,\cdot)\|^2_{L^2([0,1])}\,\big]+C_2\alpha^2\beta^2\tau\\
&+(C_5+C_2^2)\alpha^2\zeta^2\int_0^\tau\|\Theta(s,\cdot)\|^2_{L^2([0,1])}\,\mathrm{d}s\Bigg)\times\exp\left(4C_2^2\alpha^2\tau^2\right).
\end{aligned}
$$

最后，令 $N\to\infty$ 并根据单调收敛定理，我们证明了（24）。

存在性： 与之前一样，我们使用皮卡迭代论证来证明存在性。令 $U_0(s,\ x)=U(0,\ x)$ 对所有 $s\in[0,\tau]$ 都成立。皮卡迭代定义如下：

$$
\begin{aligned}
U_k(s,x)=U(0,x)&-\alpha\int_0^s\int_0^1A(x,y)U_{k-1}(u,y)\mathrm{d}y\mathrm{d}u\\
&+\alpha\beta\int_0^s\mathrm{d}\xi_2(u,x)+\alpha\zeta\int_0^s\mathrm{d}\xi_3(u,x).
\end{aligned}
\tag{52}
$$

采用与稳定性部分类似的论证，对于任意 $K\in\mathbb{N}$，我们得到：

$$
\begin{aligned}
&\mathbb{E}\big[\max_{1\le k\le K}\|U_k(s,\cdot)\|^2_{L^2([0,1])}\,\big]\le 4C_2^2\alpha^2 s\int_0^s\max_{1\le k\le K}\mathbb{E}\big[\,\|U_k(u,\cdot)\|^2_{L^2([0,1])}\,\big]\mathrm{d}u\\
&+4\Big(\mathbb{E}\big[\,\|U(0,\cdot)\|^2_{L^2([0,1])}\,\big]+C_2\alpha^2\beta^2 s+(C_5+C_2^2)\alpha^2\zeta^2\int_0^s\|\Theta(u,\cdot)\|^2_{L^2([0,1])}\,\mathrm{d}u\Big).
\end{aligned}
$$

格朗沃尔不等式意味着，对于任意 $k\in\mathbb{N}$,

$$
\begin{aligned}
&\mathbb{E}\big[\,\|U_k(s,\cdot)\|^2_{L^2([0,1])}\,\big]\\
\le\ & 4\Bigg(\mathbb{E}\big[\,\|U(0,\cdot)\|^2_{L^2([0,1])}\,\big]+C_2\alpha^2\beta^2 s+(C_5+C_2^2)\alpha^2\zeta^2\int_0^s\|\Theta(u,\cdot)\|^2_{L^2([0,1])}\,\mathrm{d}u\Bigg)\\
&\times\exp\left(4C_2^2\alpha^2 s^2\right).
\end{aligned}
$$

因此，对于任意 k和 $s\in[0,\tau]$, $U(s,\cdot)\in L^2([0,1])$. 接下来我们证明 $\{U_k\}_{k=1}^\infty$在空间 $C([0,\tau];\,L^2([0,1]))$ 中收敛。由 (52)，对于任意 $k\ge1$和 $s\in[0,\tau]$,

$$
\|U_{k+1}(s,\cdot)-U_k(s,\cdot)\|^2_{L^2([0,1])}
$$

$$
\begin{aligned}
&= \alpha^2 \left\| \int_0^s \int_0^1 A(x,y) \left(U_k(u,y) - U_{k-1}(u,y) \right) \mathrm{d}y \mathrm{d}u \right\|^2_{L^2([0,1])} \\
&\leq C_2^2 \alpha^2 s \int_0^s \| U_k(u,\cdot) - U_{k-1}(u,\cdot) \|^2_{L^2([0,1])} \, \mathrm{d}u,
\end{aligned}
$$

and

$$
\begin{aligned}
&\| U_1(s,\cdot) - U_0(s,\cdot) \|^2_{L^2([0,1])} \\
\leq & 3\alpha^2 \int_0^1 \Big(\int_0^s \int_0^1 A(x,y) U_0(y) \mathrm{d}y \mathrm{d}u \Big)^2 \mathrm{d}x + 3\alpha^2 \beta^2 \int_0^1 \Big(\int_0^s \mathrm{d}\xi_2(u,x) \Big)^2 \mathrm{d}x \\
& + 3\alpha^2 \zeta^2 \int_0^1 \Big(\int_0^s \mathrm{d}\xi_3(u,x) \Big)^2 \mathrm{d}x.
\end{aligned}
$$

因此，对于任意 $s \in [0,\tau]$，我们有

$$
\begin{aligned}
\mathbb{E}\big[\sup_{r\in[0,s]} \| U_1(r,\cdot) - U_0(r,\cdot) \|^2_{L^2([0,1])} \big] &\leq 3C_2^2 \alpha^2 s^2 \mathbb{E}\big[\| U(0,\cdot) \|^2_{L^2([0,1])} \big] + 3C_2 \alpha^2 \beta^2 s \\
&\quad + 3\left(C_5 + C_2^2 \right) \alpha^2 \zeta^2 \int_0^s \| \Theta(u,\cdot) \|^2_{L^2([0,1])} \, \mathrm{d}u \\
&:= C(s) \leq C(\tau) < \infty,
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\big[\sup_{r\in[0,s]} \| U_{k+1}(r,\cdot) - U_k(r,\cdot) \|^2_{L^2([0,1])} \big] \\
\leq & C_2^2 \alpha^2 s \int_0^s \mathbb{E}\big[\sup_{u\in[0,r]} \| U_k(u,\cdot) - U_{k-1}(u,\cdot) \|^2_{L^2([0,1])} \big] \mathrm{d}u.
\end{aligned}
$$

因此，通过归纳可得，对于任意 $k \geq 1$,

$$
\mathbb{E}\big[\sup_{r\in[0,s]} \| U_{k+1}(r,\cdot) - U_k(r,\cdot) \|^2_{L^2([0,1])} \big] \leq \frac{C(\tau) \left(C_2^2 \alpha^2 \tau s \right)^k}{k!}.
$$

因此，根据马尔可夫不等式,

$$
\mathbb{P}\big(\sup_{r\in[0,s]} \| U_{k+1}(r,\cdot) - U_k(r,\cdot) \|^2_{L^2([0,1])} > 2^{-k} \big) \leq \frac{C(\tau) \left(2C_2^2 \alpha^2 \tau s \right)^k}{k!}.
$$

设 Ω 为高斯场过程 ξ_2 和 ξ_3 所定义的路径空间。根据 Borel-Cantelli 引理，对于几乎所有 $\omega \in \Omega$，存在 $k(\omega)$，使得对于任意 $k \geq k(\omega)$，$\sup_{r\in[0,s]} \| U_{k+1}(r,\cdot) - U_k(r,\cdot) \|^2_{L^2([0,1])} \leq \frac{1}{2^k}$。因此，以概率 1，$\{U_k(\cdot,\cdot)\}_{k\geq 1}$ 在 $C([0,\tau]; L^2([0,1]))$ 中收敛，其极限为 $U(\cdot,\cdot) \in C([0,\tau]; L^2([0,1]))$。此外，通过对 (52) 取极限，可以验证 U 满足随机微分方程 (14)。因此，方程 (14) 的解存在。

唯一性： 假设存在两个解， U，U 满足随机微分方程（14）¯，由（51）可得，对任意 $s \in [0,\tau]$, $x \in [0,1]$,

$$
\overline{U}(s,x) - U(s,x) = -\alpha \int_0^s \int_0^1 A(x,y) \left(\overline{U}(u,y) - U(u,y) \right) \mathrm{d}y \mathrm{d}u,
$$

这意味着

$$\begin{aligned}&\mathbb{E}\big[\sup_{s\in[0,\tau]}\big\|\overline{U}(s,\cdot)-U(s,\cdot)\big\|_{L^2([0,1])}\big]\\ \leq& C_2^2\alpha^2\tau\mathbb{E}\big[\int_0^\tau\sup_{r\in[0,s]}\big\|\overline{U}(r,\cdot)-U(r,\cdot)\big\|_{L^2([0,1])}\mathrm{d}s\big].\end{aligned}$$

由格朗沃尔不等式可得

$$\begin{aligned}&\mathbb{E}\big[\sup_{s\in[0,\tau]}\big\|\overline{U}(s,\cdot)-U(s,\cdot)\big\|_{L^2([0,1])}\big]\\ \leq& \mathbb{E}\big[\big\|\overline{U}(0,\cdot)-U(0,\cdot)\big\|^2_{L^2([0,1])}\big]\exp\left(C_2^2\alpha^2\tau^2\right)=0.\end{aligned}$$

因此，这是 14 在 $C([0,\tau];L^2([0,1])$ 中的唯一解。 ■

7 第 2 节应用的证明2

命题 *2.1* 的证明。 首先，根据均方误差的定义$^{d,\ T}$，可得

$$\begin{aligned}\mathsf{MSE}^{d,T}(s)&=\frac{1}{d}\sum_{i=1}^d\overline{\Theta}^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})^2=\int_0^1\overline{\Theta}^{d,T}(s,y)^2\mathrm{d}y+o(1)\\ &\to\mathsf{MSE}(s):=\int_0^1\Theta(s,y)^2\mathrm{d}y\quad\text{as } d,T\to\infty,\end{aligned}$$

其中第二个等式由以下事实得出： $\Theta^{d,\ T}\in C([0,\tau];\overline{C}([0,1]))$。最后一步由定理 1.6 得出。接下来，对于预测误差，根据预测误差的定义$^{d,\ T}$，

$$\begin{aligned}\mathsf{PE}^{d,T}(s)&=\frac{1}{d^2}\sum_{i,j=1}^dA(\frac{i}{d},\frac{j}{d})\overline{\Theta}^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})\overline{\Theta}^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{j}{d})\\ &=\int_0^1\int_0^1A(x,y)\overline{\Theta}^{d,T}(s,x)\overline{\Theta}^{d,T}(s,y)\mathrm{d}x\mathrm{d}y+o(1)\\ &\to\mathsf{PE}(s)\quad\text{as } d,T\to\infty,\end{aligned}$$

其中第二个等式由以下事实得出： $\Theta^{d,\ T}\in C([0,\tau];\overline{C}([0,1])$ ）。最后一步由定理 1.6 得出。最后，刻画 Θ 的不同方程可直接由定理1.6 得出。 ■

命题 *2.2* 的证明。式（1）、（2）和（3）中的缩放条件分别对应定理 1.10 中的不同缩放。因此，对于每种情形，我们有 $\Theta^{d,\ T\,\mathrm{p}}\to\Theta$， $U^{d,\ T\,\mathrm{d}}\to U$，其中 Θ 是式（8）的解。 U 分别是式（14）、（15）、（16）在（1）、（2）和（3）中的解。接下来，我们将这些收敛结果应用于证明式（i）。

$$\begin{aligned}&\gamma\big(\mathsf{MSE}^{d,T}(s)-\mathsf{MSE}(s)\big)\\ =&\underbrace{-\gamma\big(\int_0^1\Theta(s,x)^2\mathrm{d}x-\frac{1}{d}\sum_{i=1}^d\Theta(s,\frac{i}{d})^2\big)}_{N_1^s}\underbrace{-\frac{\gamma}{d}\sum_{i=1}^d\big(\Theta(s,\frac{i}{d})^2-\Theta(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})^2\big)}_{N_2^s}\end{aligned}$$

$$\underbrace{-\frac{2\gamma}{d}\sum_{i=1}^{d}\Theta(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})\big(\Theta(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})-\overline{\Theta}^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})\big)}_{N_3^s}$$

$$+\underbrace{\frac{\gamma}{d}\sum_{i=1}^{d}\big(\Theta(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})-\overline{\Theta}^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})\big)^2}_{N_4^s}.$$

（a）对于 N_1^s，由于对任意 $s\in[0,\tau]$，$\Theta(s,\ \cdot)\in C^1([0,1])\cap L^\infty([0,1])$，我们有

$$\Big|\int_0^1\Theta(s,x)^2\mathrm{d}x-\frac{1}{d}\sum_{i=1}^{d}\Theta(s,\frac{i}{d})^2\Big|\le 2\left\|\Theta(s,\cdot)\right\|_\infty\left\|\partial_x\Theta(s,\cdot)\right\|_\infty d^{-1}$$

因此

- in (1), $N_1^s=O(T^{\frac{1}{2}d^{-1}})=o(1)$ because $T=o(d^2)$.
- in (2), $N_1^s=O(\sigma^{-1}T^{\frac{1}{2}})=o(1)$ because $\max(d,T^{\frac{1}{2}})\ll\sigma$.
- in (3), $N_1^s=O(\gamma d^{-1})=o(1)$ because $\gamma\ll d$.

(b) 对于 N_2^s，由于对任意 $x\in[0,1]$，$\Theta(\cdot,x)\in C^1([0,\tau])\cap L^\infty([0,\tau])$，我们有

$$\big|\Theta(s,\frac{i}{d})^2-\Theta(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})^2\big|\le 2\left\|\Theta(\cdot,x)\right\|_\infty\left\|\partial_s\Theta(\cdot,x)\right\|_\infty T^{-1}.$$

因此

- in (1), $N_2^s=O(\gamma T^{-1})=O(T^{-\frac{1}{2}})=o(1)$.
- in (2), $N_2^s=O(\gamma T^{-1})=O(\sigma^{-1}dT^{-\frac{1}{2}})=o(1)$ because $\max(d,T^{\frac{1}{2}})\ll\sigma$.
- in (3), $N_2^s=O(\gamma T^{-1})=o(dT^{-1})=o(1)$ because $d=O(T^{\frac{1}{2}})$.

(c) 对于 N_3^s，根据 $U^{d,T}$ 的定义，我们有

$$\begin{aligned}N_3^s&=-\frac{2}{d}\sum_{i=1}^{d}\Theta(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})U^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})\\&=-2\int_0^1\Theta(s,x)U(s,x)\mathrm{d}x\\&\quad+2\Big(\int_0^1\Theta(s,x)U(s,x)\mathrm{d}x-\frac{1}{d}\sum_{i=1}^{d}\Theta(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})U(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})\Big)\\&\quad+\frac{2}{d}\sum_{i=1}^{d}\Theta(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})\big(U(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})-U^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})\big).\end{aligned}$$

作为 $U\in C([0,\tau];\ C([0,1]))$、$U^{d,T}\xrightarrow{\mathrm{d}}U$和 $\Theta\in C^1([0,\tau];\ C^1([0,1]))$，我们有

$$\int_0^1\Theta(s,x)U(s,x)\mathrm{d}x-\frac{1}{d}\sum_{i=1}^{d}\Theta(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})U(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})=o(1)$$

$$\frac{2}{d}\sum_{i=1}^{d}\Theta(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})\big(U(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})-U^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})\big)=o(1).$$

因此，我们得到

$$N_3^s\xrightarrow{\mathrm{d}}-2\int_0^1\Theta(s,x)U(s,x)\mathrm{d}x.$$

(d) 对于 N_4^s，根据 $U^{d,T}$ 的定义，我们有

$$\begin{aligned}N_4^s&=-\frac{1}{\gamma d}\sum_{i=1}^{d}U^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})^2\\&=-\gamma^{-1}\int_0^1U(s,x)^2\mathrm{d}x-\gamma^{-1}\Big(\int_0^1U(s,x)^2\mathrm{d}x-\frac{1}{d}\sum_{i=1}^{d}U(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})^2\Big)\\&\quad-\gamma^{-1}d^{-1}\sum_{i=1}^{d}\big(U^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})^2-U(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})^2\big).\end{aligned}$$

根据（24）以及 $\gamma\gg1$, $\gamma^{-1}\int_0^1U(s,x)^2dx=o(1)$ 这一事实。由于 $U^{d,T}\xrightarrow{\mathrm{d}}U\in C([0,\tau];C([0,1]))$，我们有

$$\int_0^1U(s,x)^2\mathrm{d}x-\frac{1}{d}\sum_{i=1}^{d}U(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})^2=o(1)$$

$$\frac{1}{d}\sum_{i=1}^{d}\big(U^{d,T}(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})^2-U(\frac{\lfloor sT\rfloor}{T},\frac{i}{d})^2\big)=o(1).$$

因此，$N_4^s=o(1)$。

根据以上 (a)、(b)、(c) 和 (d) 各点，(i) 得证。在假设 1.2 下，(ii) 的结论也可类似证明。此处留给读者完成。■

引理 2.4 的证明。设对所有 $x\in[0,1]$，$\bar{A}(x)=a_0+\sum_{k=1}^{\infty}b_k\cos(2\pi kx)$。则对所有 $x,y\in[0,1]$，$A(x,y)=\bar{A}(|x-y|)$。要证明 A 满足假设 1.2，只需证明 $\bar{A}\in C^2([0,1])$。注意，$\bar{A}$ 可表示为以标准正交基 $\{1,\sqrt{2}\cos(2\pi kx)\ \sqrt{2}\sin(2\pi kx)\}_{k\geq1}$ 展开的傅里叶级数。在条件 (19) 下，我们有 $\bar{A}\in H^{(5+\varepsilon)/2}([0,1])$，其中 $H^{(5+\varepsilon)/2}([0,1])$ 是定义在 $[0,1]$ 上、具有直到 $(5+\varepsilon)$ 阶平方可积弱导数的索博列夫空间；关于 Sobolev 空间的详细信息，例如参见 Adams 和 Fournier (2003)。由索博列夫嵌入定理（Adams 和 Fournier，2003，第 4 章），因此有 $\bar{A}\in C^2([0,1])$，从而得到所需结果。

根据（21），我们有

$$B(x_1,x_2,x_3,x_4)=\sum_{p\in P_4^2}\prod_{(i,j)\in p}A(x_i,x_j)\tag{53}$$

由引理 2.4可知，A 有界且二阶连续可微。因此，B 连续且有界。此外，由式（53）可得

$$\big|B(x_1,x_3,x_1,x_3)+B(x_2,x_3,x_2,x_3)-2B(x_1,x_3,x_2,x_3)\big|$$

$$
\begin{aligned}
=&\left|A(x_1,x_3)\left(A(x_1,x_1)+A(x_2,x_2)-2A(x_1,x_2)\right)-2\left(A(x_1,x_3)-A(x_2,x_3)\right)^2\right| \\
\leq& C|x_1-x_2|^2,
\end{aligned}
$$

and

$$
\begin{aligned}
&\Big|B(x_1,x_1,x_1,x_1)+B(x_2,x_2,x_2,x_2)+6B(x_1,x_1,x_2,x_2) \\
&\qquad\qquad -4B(x_1,x_1,x_1,x_2)-4B(x_1,x_2,x_2,x_2)\Big| \\
=&3\Big|A(x_1,x_1)+A(x_2,x_2)-2A(x_1,x_2)\Big|^2 \leq C|x_1-x_2|^4,
\end{aligned}
$$

其中，最后一个不等式由引理 2.4得出。因此，函数 B 满足假设1.2。对于八阶矩，同理可得

$$
E(x_1,\cdots,x_8)=\sum_{p\in P_8^2}\prod_{(i,j)\in p}A(x_i,x_j) \tag{54}
$$

因此，根据引理 2.4， E 是连续且有界的。此外，根据（54），令$\tilde{x}=$（x_3， x_4， x_5， x_6），可得

$$
\begin{aligned}
&\Big|E(x_1,x_1,x_1,x_1,\tilde{x})+E(x_2,x_2,x_2,x_2,\tilde{x})+6E(x_1,x_1,x_2,x_2,\tilde{x}) \\
&\qquad -4E(x_1,x_1,x_1,x_2,\tilde{x})-4E(x_1,x_2,x_2,x_2,\tilde{x})\Big| \\
=&\Big|3(A(x_1,x_1)+A(x_2,x_2)-2A(x_1,x_2))^2\sum_{p\in P_{3-6}^2}\prod_{(i,j)\in p}A(x_i,x_j) \\
&+\sum_{\sigma(\mathbf{i})}\prod_{j=1}^{4}(A(x_1,x_{i_j})-A(x_2,x_{i_j})) \\
&+3(A(x_1,x_1)+A(x_2,x_2)-2A(x_1,x_2))\sum_{\sigma(\mathbf{i})}A(x_{i_3},x_{i_4})\prod_{j=1}^{2}\left(A(x_1,x_{i_j})-A(x_2,x_{i_j})\right)\Big| \\
\leq& C\left|x_1-x_2\right|^4,
\end{aligned}
$$

其中，我们用 P_{3-6}^2 表示 $\{3,4,5,6\}$ 中的所有配对，并用 $\sigma(\mathbf{i})$ 表示对集合 $=$（i_1， i_2， i_3， i_4）的所有排列中的四元组 $\mathbf{i}$ 求和，即。上面的最后一个不等式由引理2.4得出。因此，函数 E 满足假设1.2。 ■

命题 *2.5* 的证明。 我们先证明 (b) 部分。根据（8），

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\mathsf{MSE}(s)&=\frac{\mathrm{d}}{\mathrm{d}s}\int_0^1\Theta(s,x)^2\mathrm{d}x=-2\alpha\int_0^1\int_0^1\Theta(s,x)A(x,y)\Theta(s,y)\mathrm{d}x\mathrm{d}y \\
&=-2\alpha\sum_{i=1}^{\infty}\lambda_i\langle\Theta(s,\cdot),\phi_i\rangle_{L^2([0,1])}^2 \\
&\leq-2\alpha\lambda\mathsf{MSE}(s),
\end{aligned}
$$

其中，第二行的等号由我们关于 A 的假设得到，而最后一行的不等式则由假设 $\lambda:=\inf_i\lambda_i>0$ 得到。最后的结论由格朗沃尔不等式推出。

接下来，我们证明 (a) 部分。定义泛函 $\mathcal{F}\colon L^2([0,1]) \to \mathbb{R}$ 如下：

$$\mathcal{F}[\phi] := \int_0^1 \int_0^1 A(x,y)\phi(x)\phi(y)\mathrm{d}x\mathrm{d}y, \qquad \forall \phi \in L^2([0,1]).$$

其泛函梯度 $\frac{\delta \mathcal{F}[\phi]}{\delta \phi}\colon [0,1] \to \mathbb{R}$，以及其泛函 Hessian $\frac{\delta^2 \mathcal{F}[\phi]}{\delta \phi^2}\colon [0,1]^2 \to \mathbb{R}$，分别为

$$\frac{\delta \mathcal{F}[\phi]}{\delta \phi}(x) = 2\int_0^1 A(x,y)\phi(y)\mathrm{d}y, \quad \frac{\delta^2 \mathcal{F}[\phi]}{\delta \phi^2}(x,y) = 2A(x,y)$$

设 Θ 是方程（8）的解。对于任意 $s \in [0,\tau]$，我们有

$$\begin{aligned}\mathsf{PE}'(s) &= \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{F}[\Theta(s,\cdot)] = -2\alpha \int_0^1 \int_0^1 \int_0^1 A(x,y)A(y,z)\Theta(s,x)\Theta(s,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z \\ &= -\frac{\alpha}{2}\int_0^1 \left|\frac{\delta \mathcal{F}[\Theta(s,\cdot)]}{\delta \Theta(s,\cdot)}(y)\right|^2 \mathrm{d}y,\end{aligned}$$

其中，第二个等式由（8）可得。在区间 $[0,s]$ 上积分，得到

$$\begin{aligned}\mathsf{PE}(s) &= \mathsf{PE}(0) - \frac{\alpha}{2}\int_0^s \int_0^1 \left|\frac{\delta \mathcal{F}[\Theta(r,\cdot)]}{\delta \Theta(r,\cdot)}(y)\right|^2 \mathrm{d}y\mathrm{d}r \\ &= \mathcal{F}[\Theta(0,\cdot)] - \mathcal{F}[0] - \frac{\alpha}{2}\int_0^s \int_0^1 \left|\frac{\delta \mathcal{F}[\Theta(r,\cdot)]}{\delta \Theta(r,\cdot)}(y)\right|^2 \mathrm{d}y\mathrm{d}r, \end{aligned} \tag{55}$$

其中，最后一步由以下事实可得：$\mathcal{F}[0] = 0$。注意，

$$\frac{\delta \mathcal{F}[\Theta(r,\cdot)]}{\delta \Theta(r,\cdot)}(y) = 2\int_0^1 A(x,y)\Theta(r,y)\mathrm{d}y = -\frac{2}{\alpha}\partial_r \Theta(r,x). \tag{56}$$

因此，我们有

$$\begin{aligned}\mathsf{PE}(0) - \mathcal{F}[0] &= -\left\langle \frac{\delta \mathcal{F}[\Theta(0,\cdot)]}{\delta \Theta(0,\cdot)}, -\Theta(0,\cdot)\right\rangle_{L^2([0,1])} - \int_0^1 \int_0^1 \Theta(0,x)A(x,y)\Theta(0,y)\mathrm{d}x\mathrm{d}y \\ &\le -\frac{2}{\alpha}\langle \partial_s \Theta(s,x)|_{s=0}, \Theta(0,\cdot)\rangle_{L^2([0,1])} \\ &= -\frac{1}{\alpha}\left(\frac{\mathrm{d}}{\mathrm{d}s}\int_0^1 \Theta(s,y)^2\mathrm{d}y\right)\Big|_{s=0}, \end{aligned} \tag{57}$$

其中，第一步由泰勒展开得到，不等式则由我们关于 A 以及（56）的假设得到。同时，根据（56），

$$\begin{aligned} &-\frac{\alpha}{2}\int_0^s \int_0^1 \left|\frac{\delta \mathcal{F}[\Theta(r,\cdot)]}{\delta \Theta(r,\cdot)}(y)\right|^2 \mathrm{d}y\mathrm{d}r \\ &= -\frac{2}{\alpha}\int_0^s \int_0^1 (\partial_r \Theta(r,y))^2\,\mathrm{d}y\mathrm{d}r = -\frac{2}{\alpha}\int_0^1 \int_0^s \partial_r \Theta(r,y)\Theta(r,y)\mathrm{d}r\mathrm{d}y \\ &= -\frac{2}{\alpha}\int_0^1 \Theta(s,y)\partial_r \Theta(r,y)|_{r=s} - \Theta(0,y)\partial_r \Theta(r,y)|_{r=0}\mathrm{d}y \\ &\quad + \frac{2}{\alpha}\int_0^1 \int_0^s \Theta(r,y)\partial_{rr}^2 \Theta(r,y)\mathrm{d}r\mathrm{d}y \end{aligned} \tag{58}$$

$$= \frac{2}{\alpha}\left(\frac{\mathrm{d}}{\mathrm{d}r}\int_0^1 \Theta(r,x)^2\mathrm{d}x\Big|_{r=0} - \frac{\mathrm{d}}{\mathrm{d}r}\int_0^1 \Theta(r,x)^2\mathrm{d}x\Big|_{r=s}\right) + \frac{\alpha}{2}\int_0^s\int_0^1 \left|\frac{\delta\mathcal{F}[\Theta(r,\cdot)]}{\delta\Theta(r,\cdot)}(y)\right|^2 \mathrm{d}y\mathrm{d}r \tag{59}$$

其中，最后一步成立是因为

$$\begin{aligned}
&\frac{2}{\alpha}\int_0^1\int_0^s \Theta(r,y)\partial_{rr}^2\Theta(r,y)\mathrm{d}r\mathrm{d}y \\
&= -2\int_0^1\int_0^s \Theta(r,y)\int_0^1 A(y,z)\partial_r\Theta(r,z)\mathrm{d}z\mathrm{d}r\mathrm{d}y \\
&= 2\alpha\int_0^1\int_0^s \Theta(r,y)\int_0^1 A(y,z)\int_0^1 A(z,x)\Theta(r,x)\mathrm{d}x\mathrm{d}z\mathrm{d}r\mathrm{d}y \\
&= 2\alpha\int_0^s\int_0^1 \left(\int_0^1 A(y,z)\Theta(r,z)\mathrm{d}z\right)^2\mathrm{d}y\mathrm{d}s \\
&= \frac{\alpha}{2}\int_0^s\int_0^1 \left|\frac{\delta\mathcal{F}[\Theta(r,\cdot)]}{\delta\Theta(r,\cdot)}(y)\right|^2\mathrm{d}y\mathrm{d}r.
\end{aligned}$$

因此，结合（55）、（57）和（58），我们得到

$$\begin{aligned}
&\mathsf{PE}(s) \\
&\le -\frac{1}{\alpha}\left(\frac{\mathrm{d}}{\mathrm{d}s}\int_0^1 \Theta(s,y)^2\mathrm{d}y\right)\Big|_{s=0} + \frac{1}{\alpha}\left(\frac{\mathrm{d}}{\mathrm{d}r}\int_0^1 \Theta(r,x)^2\mathrm{d}x\Big|_{r=0} - \frac{\mathrm{d}}{\mathrm{d}r}\int_0^1 \Theta(r,x)^2\mathrm{d}x\Big|_{r=s}\right) \\
&= -\frac{1}{\alpha}\frac{\mathrm{d}}{\mathrm{d}s}\int_0^1 \Theta(s,x)^2\mathrm{d}x.
\end{aligned}$$

对 $[0,\tau]$ 积分，得到

$$\int_0^\tau \mathsf{PE}(s)\mathrm{d}s \le -\frac{\mathsf{MSE}(\tau)}{\alpha} + \frac{\mathsf{MSE}(0)}{\alpha} \le \frac{\mathsf{MSE}(0)}{\alpha}.$$

根据先前的计算，我们知道 $\mathsf{PE}'(s) = -\frac{\alpha}{2}\int_0^1 \left|\frac{\delta\mathcal{F}[\Theta(s,\cdot)]}{\delta\Theta(s,\cdot)}(y)\right|^2\mathrm{d}y \le 0$。因此

$$\mathsf{PE}(\tau) \le \frac{\int_0^\tau \mathsf{PE}(s)\mathrm{d}s}{\tau} \le \frac{\mathsf{MSE}(0)}{\alpha\tau}.$$

■

参考文献

E. Abbe、E. B. Adsera 和 T. Misiakiewicz。《合并阶梯性质：两层神经网络上随机梯度下降学习稀疏函数的必要且近乎充分条件》。载于学习理论会议，第 4782–4887 页。PMLR，2022年。（引自第12页。）

R. A. Adams 和 J. J. Fournier。*Sobolev*空间。Elsevier，2003年。（引自第 56页。）

R. J. Adler 和 J. E. Taylor。随机场与几何，第 80 卷。Springer，2007年。（引文见第4页和第48页。）

B. Agrawalla、K. Balasubramanian 和 P. Ghosal。在线最小二乘随机梯度下降的线性泛函的高维中心极限定理。arXiv 预印本 arXiv:2302.09727，2023年。（引文见第 12 页。）A. Anastasiou、K. Balasubramanian 和 M. A. Erdogdu。基于鞅中心极限定理非渐近速率的随机梯度下降正态近似。载于学习理论会议。PMLR，2019年。（引文见第 13 页。）V. I. Arnold。《常微分方程》。施普林格科学与商业媒体，1992年。（引文见第 3、20、31 和 44 页。）H. Asi 和 J. C. Duchi。随机（近似）近端点方法：收敛、最优性与自适应性。《SIAM 优化杂志》，2019年。（引文见第 13 页。）J. Ba、M. A. Erdogdu、T. Suzuki、Z. Wang、D. Wu 和 G. Yang。特征学习的高维渐近：梯度步如何改进表示。载于《神经信息处理系统进展》，2022年。（引文见第 12 页。）M. Bal´azs 和 T. Sepp¨al¨ainen。非对称简单排斥过程中的电流方差阶数与扩散率。《数学年刊》（第 2 辑），171(2):1237–1265，2010年。（引文见第 13 页。）A. Barakat、P. Bianchi、W. Hachem 和 S. Schechtman。带动量的随机优化：收敛、涨落与陷阱规避。《电子统计学杂志》，15(2):3892–3947，2021年。（引文见第 13 页。）P. L. Bartlett、P. M. Long、G. Lugosi 和 A. Tsigler。线性回归中的良性过拟合。《美国国家科学院院刊》，117(48):30063–30070，2020年。（引文见第 7 页。）G. Ben Arous、R. Gheissari 和 A. Jagannath。来自高维推断的非凸损失在线随机梯度下降。《机器学习研究杂志》，22(1):4788–4838，2021年。（引文见第 12 页。）G. Ben Arous、R. Gheissari 和 A. Jagannath。随机梯度下降的高维极限定理：有效动力学与临界标度。《神经信息处理系统进展》，35:25349–25362，2022年。（引文见第 11 页。）A. Benveniste、M. M´etivier 和 P. Priouret。《自适应算法与随机逼近》，第 22 卷。施普林格科学与商业媒体，2012年。（引文见第 12 页。）L. Bertini 和 G. Giacomin。来自粒子系统的随机 Burgers 方程与 KPZ 方程。《数学物理通讯》，183(3):571–607，1997年。（引文见第 13 页。）P. Billingsley。概率测度的收敛。John Wiley & Sons，2009。（引文见第 17、28 和 40 页。）L. Bottou、F. E. Curtis 和 J. Nocedal。大规模机器学习的优化方法。《SIAM 评论》，60(2):223–311，2018年。（引文见第 13 页。）A. Caponnetto 和 E. De Vito。正则化最小二乘算法的最优速率。《计算数学基础》，7:331–368，2007年。（引文见第 7 页。）M. Celentano、C. Cheng 和 A. Montanari。随机数据下的一阶方法的高维渐近分析。arXiv 预印本 arXiv:2112.07572，2021年。（引文见第 12 页。）

K. A. Chandrasekher、M. Lou 和 A. Pananjady。广义秩一矩阵感知的交替最小化：随机初始化下的精确预测。arXiv 预印本 arXiv:2207.09660，2022年。（引自第 12 页。）K. A. Chandrasekher、A. Pananjady 和 C. Thrampoulidis。随机数据下迭代非凸优化的精确全局收敛保证。《统计年鉴》，51(1):179–210，2023年。（引自第 12 页。）H. Chen、W. Lu 和 R. Song。通过随机梯度下降进行在线决策的统计推断。《美国统计学会会刊》，116(534):708–719，2021年。（引自第 8 页。）L. Chizat 和 F. Bach。利用最优传输研究过参数化模型中梯度下降的全局收敛性。神经信息处理系统进展，31，2018年。（引自第 12 页。）I. Corwin。Kardar–Parisi–Zhang 方程与普适类。《随机矩阵：理论与应用》，1(01):1130001，2012年。（引自第 13 页。）I. Corwin、H. Shen 和 L.-C. Tsai。ASEP(q, j) 收敛到 KPZ 方程。《亨利·庞加莱研究所概率与统计年鉴》，54(2):995–1012，2018年。（引自第 13 页。）I. Corwin、P. Ghosal、H. Shen 和 L.-C. Tsai。六顶点模型的随机偏微分方程极限。《数学物理通讯》，375(3):1945–2038，2020年。（引自第 13 页。）G. Da Prato 和 J. Zabczyk。无限维空间中的随机方程。剑桥大学出版社，2014年。（引自第 3 页。）R. Darling 和 J. Norris。马尔可夫链的微分方程近似。《概率调查》，5:37–79，2008年。（引自第 3 页。） D. Dauvergne 、 J. Ortmann 和 B. Vir´ag。有向景观。《数学学报》，229(2):201–285，2022 年。（引自第 13 页。）D. Davis、D. Drusvyatskiy 和 L. Jiang。非光滑随机逼近中的渐近正态性与最优性。arXiv 预印本 arXiv:2301.06632，2023年。（引自第 13 页。）P. Deift 和 T. Trogdon。随机数据数值计算中的普适性：案例研究与解析结果。《数学物理学杂志》，60(10):103306，2019年。（引自第 12 页。）A. Dieuleveut、N. Flammarion 和 F. Bach。更难、更好、更快、更强：最小二乘回归的收敛速率。《机器学习研究杂志》，18(101):1–51，2017年。（引自第 7 页。）A. Dieuleveut、A. Durmus 和 F. Bach。弥合固定步长随机梯度下降与马尔可夫链之间的差距。《统计年鉴》，48(3):1348–1382，2020年。（引自第 13 页。）X. Ding 和 T. Trogdon。一般尖峰协方差矩阵类上的共轭梯度算法。《应用数学季刊》，80(1):99–155，2022年。（引自第 12 页。）Z. Ding、S. Chen、Q. Li 和 S. J. Wright。深度 ResNet 的过参数化：零损失与平均场分析。《机器学习研究杂志》，2022年。（引自第 12 页。）J. Duchi 和 F. Ruan。随机优化中的渐近最优性。《统计年鉴》，2020年。（引自第 13 页。）

A. Durmus、E. Moulines、A. Naumov、S. Samsonov、K. Scaman 和 H.-T. Wai。固定步长线性随机逼近的紧高概率界。神经信息处理系统进展，34:30063–30074，2021年。（引自第 12 页。）A. Durmus、E. Moulines、A. Naumov 和 S. Samsonov。线性随机逼近的 Polyak-Ruppert平均迭代的有限时间高概率界。arXiv 预印本 arXiv:2207.04475，2022年。（引自第 12 页。）Y. Fang、J. Xu 和 L. Yang。随机梯度下降估计量的在线自助法置信区间。《机器学习研究杂志》，2018年。（引自第 11 页。）X. Fontaine、V. De Bortoli 和 A. Durmus。随机梯度下降及其连续时间对应过程的收敛速率与逼近结果。载于《学习理论会议》，第 1965–2058 页。PMLR，2021年。（引自第 13 页。）C. Gerbelot、E. Troiani、F. Mignacco、F. Krzakala 和 L. Zdeborova。随机梯度下降方法的严格动力学平均场理论。arXiv 预印本 arXiv:2210.06591，2022年。（引自第 12 页。）B. Gess、R. S. Gvalani 和 V. Konarovskyi。作为随机梯度下降涨落平均场极限的守恒型随机偏微分方程。arXiv 预印本 arXiv:2207.05705，2022年。（引自第 12 页。）P. Ghosal。Hall-Littlewood-PushTASEP 及其 KPZ 极限。arXiv 电子预印本，文章编号 arXiv:1701.07308，2017年1月。doi: 10.48550/arXiv.1701.07308。（引自第 13 页。）R. M. Gray 等。Toeplitz 矩阵与循环矩阵：综述。《通信信息论基础与趋势®》，2(3):155–239，2006年。（引自第 3 页。）A. Gupta、H. Chen、J. Pi 和 G. Tendolkar。递归随机算法诱导的马尔可夫链的一些极限性质。SIAM数据科学数学杂志，2(4):967–1003，2020年。（引自第 13 页。）G. Kallianpur 和 J. Xiong。无限维空间中的随机微分方程。讲义-专著丛书，26:iii–342，1995年。（引自第 3 页。）I. Karatzas 和 S. Shreve。《布朗运动与随机微积分》，第 113 卷。施普林格科学与商业媒体，2012年。（引自第 26 页。）K. Khamaru、Y. Deshpande、T. Lattimore、L. Mackey 和 M. J. Wainwright。自适应线性回归中的近最优推断。arXiv 预印本 arXiv:2107.02266v3，2021年。（引自第 8 页。）C. Kipnis 和 C. Landim。相互作用粒子系统的尺度极限，第 320 卷。施普林格科学与商业媒体，1998年。（引自第 3 页。）W. Krichene 和 P. L. Bartlett。随机下降动力学中的加速与平均。神经信息处理系统进展，30，2017年。（引自第 12 页。）H. J. Kushner 和 D. S. Clark。约束与无约束系统的随机逼近方法，第 26 卷。施普林格科学与商业媒体，2012年。（引自第 12 页。）H. J. Kushner 和 G. George Yin。随机逼近、递归算法及其应用，第 35 卷。施普林格科学与商业媒体，2003年。（引自第 12 页。）

G. Lan. 面向机器学习的一阶与随机优化方法，第1卷。Springer，2020年。（引见第13页。）C. J. Li、M. Wang、H. Liu 和 T. Zhang. 在线主成分估计的近最优随机逼近。《数学规划》，167(1):75–97，2018年。（引见第13页。）Q. Li、C. Tai 和 E. Weinan. 随机修正方程与随机梯度算法的动力学 I：数学基础。机器学习研究杂志，20(1):1474–1520，2019年。（引见第12页。）T. Liggett. 随机相互作用系统：接触过程、选民过程与排斥过程，第324卷。施普林格科学与商业媒体，1999年。（引见第3页。）J. Lin 和 L. Rosasco. 多遍随机梯度方法的最优速率。《机器学习研究杂志》，18(97):1–47，2017年。（引见第7页。）J. Liu 和 Y. Yuan. 论随机梯度方法的几乎必然收敛速率。载于学习理论会议，第2963–2983页。PMLR，2022年。（引见第13页。）L. Ljung、G. Pflug 和 H. Walk. 随机逼近与随机系统优化，第17卷。Birkh¨auser，2012年。（引见第12页。）G. J. Lord、C. E. Powell 和 T. Shardlow. 计算随机偏微分方程导论，第50卷。剑桥大学出版社，2014年。（引见第17页和第20页。）K. Matetski、J. Quastel 和 D. Remenik. KPZ 固定点。《数学学报》，227(1):115–203，2021年。（引见第13页。）J. C. Mattingly、N. S. Pillai 和 A. M. Stuart. 高维随机游走Metropolis算法的扩散极限。《应用概率年刊》，22(3):881–930，2012年。（引见第13页。）S. Mei、A. Montanari 和 P.-M. Nguyen. 两层神经网络损失景观的平均场视角。《美国国家科学院院刊》，115(33):E7665–E7671，2018年。（引见第12页。）P. Mertikopoulos、N. Hallak、A. Kavis 和 V. Cevher. 论非凸问题中随机梯度下降的几乎必然收敛。神经信息处理系统进展，33:1117–1128，2020年。（引见第13页。）S. Meyn. 控制系统与强化学习。剑桥大学出版社，2022年。（引见第8页。）N. M¨ucke、G. Neu 和 L. Rosasco. 利用尾部平均和小批量处理突破随机梯度下降饱和。神经信息处理系统进展，32，2019年。（引见第7页。）S. Mukherjee 和 S. Sen. 高维线性回归中的变分推断。《机器学习研究杂志》，23(304):1–56，2022年。（引见第13页。）C. Paquette 和 E. Paquette. 大规模二次模型上随机动量方法的动力学。神经信息处理系统进展，34:9229–9240，2021年。（引见第12页。）

C. Paquette、K. Lee、F. Pedregosa 和 E. Paquette. 大规模随机梯度下降：平均情形分析、渐近分析与步长临界性。载于学习理论会议，第3548–3626页。PMLR，2021年。（引见第12页。）C. Paquette、E. Paquette、B. Adlam 和 J. Pennington. 高维随机梯度下降的均质化：精确动力学与泛化性质。arXiv 预印本 arXiv:2205.07069，2022a。（引见第12页。）C. Paquette、E. Paquette、B. Adlam 和 J. Pennington. 隐式正则化还是隐式条件化？高维随机梯度下降的精确风险轨迹。载于神经信息处理系统进展，2022b。（引见第12页。）C. Paquette、B. van Merri¨enboer、E. Paquette 和 F. Pedregosa. 大型模型的停止时间是可预测的：一种普适性与平均情形分析。计算数学基础，第1–77页，2022c。（引见第12页。）H. T. Pham 和 P.-M. Nguyen. 平均场训练下多层神经网络的极限涨落与轨迹稳定性。神经信息处理系统进展，34:4843–4855，2021年。（引见第12页。）N. S. Pillai、A. M. Stuart 和 A. H. Thi´ery. 高维朗之万算法的最优缩放与扩散极限。《应用概率年刊》，第2320–2356页，2012年。（引见第13页。）L. Pillaud-Vivien、A. Rudi 和 F. Bach. 通过多次遍历实现随机梯度下降在困难学习问题上的统计最优性。神经信息处理系统进展，31，2018年。（引见第7页。）B. T. Polyak 和 A. B. Juditsky. 通过平均加速随机逼近。SIAM控制与优化杂志，30(4):838–855，1992年。（引见第13页。）J. Qiu 和 S. Sen. 高维线性回归的TAP自由能。arXiv 预印本 arXiv:2203.07539，2022年。（引见第13页。）J. Quastel. KPZ 导论。数学当前进展，2011年(1)，2011年。（引见第13页。）J. Quastel 和 S. Sarkar. 排斥过程与 KPZ 方程向 KPZ 固定点的收敛。《美国数学会杂志》，36(1):251–289，2023年。（引见第13页。）D. Revuz 和 M. Yor. 连续鞅与布朗运动，数学科学基础丛书 ⌈数学科学基本原理⌋第293卷。Springer-Verlag，柏林，第3版，1999年。ISBN 3-540-64325-7。DOI：10.1007/978-3-662-06400-9。网址：https://doi.org/10.1007/978-3-662-06400-9。（引见第51页。）H. Robbins 和 S. Monro. 一种随机逼近方法。数学统计年刊，第400–407页，1951年。（引见第2页。）G. Rotskoff 和 E. Vanden-Eijnden. 人工神经网络的可训练性与精度：一种相互作用粒子系统方法。纯粹与应用数学通讯，75(9):1889–1935，2022年。（引见第12页。）

D. Ruppert. 对缓慢收敛的Robbins-Monro过程进行高效估计。技术报告，康奈尔大学运筹学与工业工程系，1988年。（引自第13页。）O. Sebbouh、R. M. Gower和A. Defazio。随机梯度下降和随机重球法的几乎必然收敛速率。载于《学习理论会议》，第3935–3971页。PMLR，2021年。（引自第13页。）Q.-M. Shao和Z.-S. Zhang。多元非线性统计量的Berry–Esseen界及其在M估计量和随机梯度下降算法中的应用。伯努利，28(3):1548–1576，2022年。（引自第13页。）J. Sirignano和K. Spiliopoulos。神经网络的平均场分析：大数定律。SIAM应用数学杂志，80(2):725–752，2020a。（引自第12页。）J. Sirignano和K. Spiliopoulos。神经网络的平均场分析：中心极限定理。随机过程及其应用，130(3):1820–1852，2020b。（引自第12页。）J. Sirignano和K. Spiliopoulos。深度神经网络的平均场分析。运筹学数学，47(1):120–152，2022年。（引自第12页。）H. Spohn。相互作用粒子的大尺度动力学。施普林格科学与商业媒体，2012年。（引自第3页。）A.-S. Sznitman。随机环境中的随机游走专题。载于《概率论学校与会议：2002年5月13–17日》，第17卷，第203–266页，2004年。（引自第3页。）Y. S. Tan和R. Vershynin。任意初始化下的在线随机梯度下降可解决非光滑非凸相位恢复问题。《机器学习研究杂志》，24，2023年。（引自第12页。）P. Toulis和E. M. Airoldi。基于随机梯度的估计量的渐近性质与有限样本性质。《统计年鉴》，45(4):1694–1727，2017年。（引自第13页。）M. Udell和A. Townsend。为什么大数据矩阵近似为低秩？SIAM数据科学数学杂志，1(1):144–160，2019年。（引自第7页。）R. Veiga、L. Stephan、B. Loureiro、F. Krzakala和L. Zdeborov´a。高维双层神经网络中随机梯度下降的相图。载于《神经信息处理系统进展》，2022年。（引自第11页。）C. Vignat。高斯随机向量位置混合的广义Isserlis定理。统计与概率快报，82(1):67–71，2012年。（引自第16页。）C. Wang和Y. M. Lu。高维在线独立成分分析的标度极限。统计力学杂志，2019(12)，2019年。（引自第11页。）C. Wang、J. Mattingly和Y. M. Lu。标度极限：在线学习算法的精确且易处理分析及其在正则化回归和主成分分析中的应用。arXiv预印本 arXiv:1712.04332，2017年。（引自第11页。）L. Yu、K. Balasubramanian、S. Volgushev和M. A. Erdogdu。非凸情形下固定步长随机梯度下降的分析：渐近正态性与偏差。《神经信息处理系统进展

g，2021 （引用见第 13 页。）
年。

O. Zeitouni。《随机环境中的随机游走》。 数学讲义，1837:190–312，2004年。 （引用见第 3 页。）
B. Zuo、C. Yin 和 N. Balakrishnan。椭圆分布联合矩的表达式。计算与应用数学杂志，391:113418，2021年。 （引用见第 16 页。）